GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 6, 2004, 13:10:02 ; Search time 49 Seconds
(without alignments)
5634.254 Million cell updates/sec

Title: US-10-005-480A-743
Perfect score: 4804
Sequence: 1 MESTLTLATEQPVKKNTLKK.....QPVSEILQLKTYLPTFETTI 875

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2540.5	52.9	565	11	Q99K07	Q99k07 mus musculu
2	2435	50.7	906	11	Q924C4	Q924c4 mus musculu
3	2191	45.6	874	13	Q7ZXN7	Q7zxn7 xenopus lae
4	1360	28.3	251	4	Q9NQM9	Q9nqm9 homo sapien
5	1263	26.3	817	12	Q9J5H1	Q9j5h1 fowlpox vir
6	1263	26.3	817	12	O90761	O90761 fowlpox vir
7	972	20.2	415	11	Q8CAF0	Q8caf0 mus musculu
8	837	17.4	195	11	Q91VZ7	Q91vz7 mus musculu
9	835	17.4	152	4	Q9H515	Q9h515 homo sapien
10	755.5	15.7	479	10	Q9FS13	Q9fs13 spinacia ol
11	746	15.5	274	4	Q7Z3P5	Q7z3p5 homo sapien
12	743.5	15.5	496	10	Q9SU83	Q9su83 arabidopsis
13	737.5	15.4	479	10	Q42974	Q42974 oryza sativ
14	713.5	14.9	457	10	Q84WJ3	Q84wj3 arabidopsis
15	712.5	14.8	457	10	Q9SU82	Q9su82 arabidopsis
16	680.5	14.2	461	10	Q9SU81	Q9su81 arabidopsis
17	677.5	14.1	151	11	Q9QYV2	Q9qyv2 rattus norv
18	671	14.0	257	11	O88827	O88827 rattus norv
19	644	13.4	477	11	Q921P7	Q921p7 mus musculu
20	644	13.4	477	11	Q9EQG7	Q9eqg7 mus musculu
21	637	13.3	485	3	O94323	O94323 schizosacch
22	635	13.2	757	5	P90755	P90755 caenorhabdi
23	632.5	13.2	477	4	Q9UJA9	Q9uja9 homo sapien
24	630.5	13.1	429	2	O69013	O69013 zymomonas m
25	629.5	13.1	453	4	Q9Y6X5	Q9y6x5 homo sapien
26	623	13.0	829	5	P90754	P90754 caenorhabdi
27	607.5	12.6	456	11	Q8BTJ4	Q8btj4 mus musculu
28	604.5	12.6	453	11	Q8K1L3	Q8k1l3 mus musculu
29	581.5	12.1	281	6	Q9TSB2	Q9tsb2 bos taurus
30	578.5	12.0	432	16	Q8PIS1	Q8pis1 xanthomonas
31	553	11.5	451	16	Q9A8S4	Q9a8s4 caulobacter
32	548.5	11.4	424	16	Q8P7F7	Q8p7f7 xanthomonas
33	518.5	10.8	440	4	Q96M57	Q96m57 homo sapien
34	510	10.6	440	11	Q8BGN3	Q8bgn3 mus musculu
35	510	10.6	458	4	Q8IUS8	Q8ius8 homo sapien
36	502.5	10.5	433	16	Q87A55	Q87a55 xylella fas
37	499.5	10.4	433	16	Q9PAB9	Q9pab9 xylella fas
38	356.5	7.4	234	10	Q94K08	Q94k08 arabidopsis
39	353.5	7.4	497	5	Q17540	Q17540 caenorhabdi
40	278	5.8	465	5	Q22129	Q22129 caenorhabdi
41	244	5.1	51	4	Q9NTT8	Q9ntt8 homo sapien
42	222.5	4.6	434	16	Q97LS7	Q97ls7 clostridium
43	221.5	4.6	431	16	Q8XKL9	Q8xkl9 clostridium
44	212	4.4	133	10	Q9SU80	Q9su80 arabidopsis
45	209	4.4	422	16	Q8Y778	Q8y778 listeria mo

ALIGNMENTS

RESULT 1
Q99K07
ID Q99K07 PRELIMINARY; PRT; 565 AA.
AC Q99K07;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to ectonucleotide pyrophosphatase/phosphodiesterase 3
DE (Fragment).
GN ENPP3 OR AI876438.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC005527; AAH05527.1; -.
DR MGD; MGI:2143702; Enpp3.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0009117; P:nucleotide metabolism; IEA.
DR InterPro; IPR001604; Endonuclease.
DR InterPro; IPR002591; Phosphodiest.
DR Pfam; PF01223; Endonuclease; 1.
DR Pfam; PF01663; Phosphodiest; 1.
DR SMART; SM00477; NUC; 1.
FT NON_TER 1 1
SQ SEQUENCE 565 AA; 63895 MW; 7179104F18634A04 CRC64;

Query Match 52.9%; Score 2540.5; DB 11; Length 565;
Best Local Similarity 81.1%; Pred. No. 3.1e-207;
Matches 460; Conservative 53; Mismatches 51; Indels 3; Gaps 3;

```
Qy        310 KAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLHNCVNIIL 369
              ||:|| |||:| |||||:||: ||||| |||||| ||:||||||||||||||||||||:
Db          1 KADRPSFYTIYVEEPDSAGHSSGPVSAGVIKALQSVDNAFGMLMEGLKQRNLHNCVNIIV 60
```

```
Qy        370 LADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNSEEIVRNLSC 429
Db         61 LADHGMDQTSCDRVEYMTDYFPKIN-FYMYQGPAPRIRTRNIPQDFFTFNSEEIVRNLSC 119

Qy        430 RKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGGGNHGYNNEF 489
Db        120 RKPDQHFKPYLTPDLPKRLHYAKNVRIDKAHLMVDRQWLAFRSKGSSNCGGGTHGYNNEF 179

Qy        490 RSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNHLLKVPFYE 549
Db        180 KSMEAIFLAHGPSFIEKTVIEPFENIEVYNLLCDLLHIEPAPNNGTHGSLNHLLKTPFYK 239

Qy        550 PSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQL-EQVNQMLNLTQEEITATVKVNL 608
Db        240 PSHAGELSTPAVCGFTTPLPTDPLDCSCPALQNTPGLEEQANQRLNLSEGEVAATVKANL 299

Qy        609 PFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPTVPDCLRAD 668
Db        300 PFGRPRVMQKNGDHCLLYHRDYISGYGKAMKMPMWSSYTVLKPGDTSSLPPTVPDCLRAD 359

Qy        669 VRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEEFRKMWDYF 728
Db        360 VRVAPSESQKCSFYLADKNITHGFLY-PAIKGTNESRYDALITSNLVPMYKEFKKMWDYF 418

Qy        729 HSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHYFVVLTSCK 788
Db        419 HEVLLIKYAIERNGLNVVSGPIFDYNYDGHFDAPDEITQYVAGTDVPIPTHYFVVLTSCK 478

Qy        789 NKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVRDVELLTGL 848
Db        479 DQTHTPDSCPGWLDVLPFIVPHRPTNIESCSENKTEDLWVEERFQAHVARVRDVELLTGL 538

Qy        849 DFYQDKVQPVSEILQLKTYLPTFETTI 875
Db        539 DFYQEKAQPVSQILQLKTYLPTFETII 565
```

RESULT 2
Q924C4
ID Q924C4 PRELIMINARY; PRT; 906 AA.
AC Q924C4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ectonucleotide pyrophosphatase/phosphodiesterase 1 allotype b
DE (EC 3.1.4.1).
GN ENPP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Banakh I., Sali A., Dubljevic V., Grobben B., Slegers H., Goding J.W.;
RT "Structural basis of allotypes of ecto-nucleotide
RT pyrophosphatase/phosphodiesterase (plasma cell membrane glycoprotein
RT PC-1) in the mouse and rat.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF339910; AAK84174.1; -.
DR MGD; MGI:97370; Enpp1.
DR GO; GO:0001501; P:skeletal development; IMP.
DR InterPro; IPR001604; Endonuclease.
DR InterPro; IPR002591; Phosphodiest.
DR InterPro; IPR001212; Somatomedin_B.
DR Pfam; PF01223; Endonuclease; 1.
DR Pfam; PF01663; Phosphodiest; 1.
DR Pfam; PF01033; Somatomedin_B; 2.
DR PRINTS; PR00022; SOMATOMEDINB.
DR SMART; SM00477; NUC; 1.
DR SMART; SM00201; SO; 2.
DR PROSITE; PS00524; SOMATOMEDIN_B; 2.
KW Hydrolase.

SQ SEQUENCE 906 AA; 103125 MW; 1FA071071E4B9C50 CRC64;

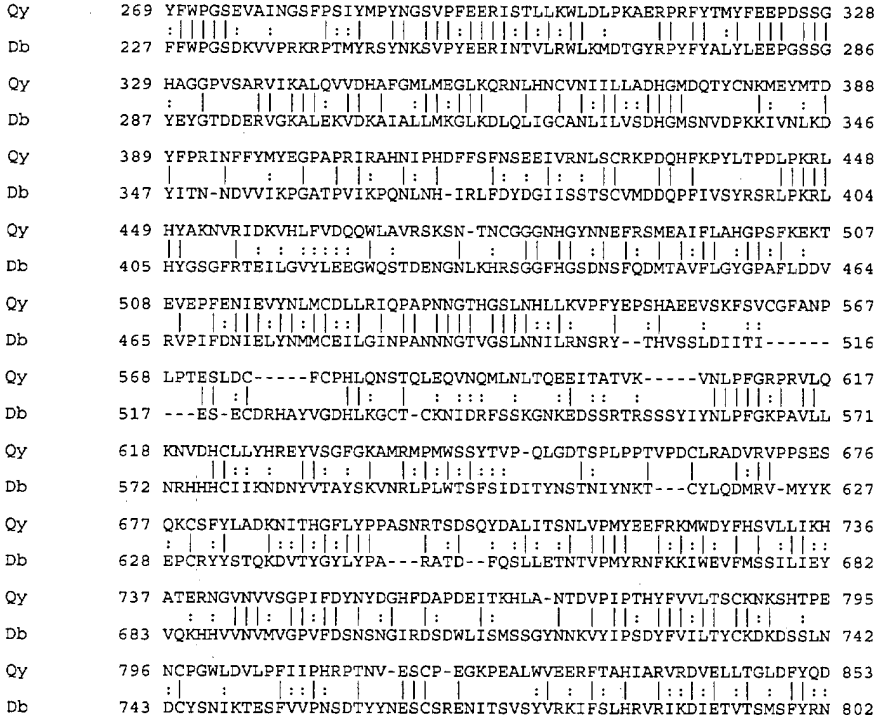
Query Match 50.7%; Score 2435; DB 11; Length 906;
Best Local Similarity 51.0%; Pred. No. 5.8e-198;
Matches 447; Conservative 152; Mismatches 250; Indels 28; Gaps 9;

```
Qy         10 EQPVKK----------NTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCF 59
Db         39 EEPLEKAERARPAKDPNT---YKVLSLVLSVCVLTTILGCIFGLKPSCAKEVKSCKGRCF 95

Qy         60 DASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDC 119
Db         96 ERTF---SNCRCDAACVSLGNCCLDFQETCVEPTHIWTCNKFRCGEKRLSRFVCSCADDC 152

Qy        120 LQKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDT 179
Db        153 KTHNDCCINYSSVCQDKKSWVEETCESIDTPECPAEFESPPTLLFSLDGFRAEYLHTWGG 212

Qy        180 LMPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSL 239
Db        213 LLPVISKLKNCGTYTKNMRPMYPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMNASFSL 272

Qy        240 SSKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERI 299
Db        273 KSKEKFNPLWYKGQPIWVTANHQEVKSGTYFWPGSDVEIDGILPDIYKVYNGSVPFEERI 332

Qy        300 STLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQR 359
Db        333 LAVLEWLQLPSHERPHFYTLYLEEPDSSGHSHGPVSSEVIKALQKVDRLVGMLMDGLKDL 392

Qy        360 NLHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFN 419
Db        393 GLDKCLNLILISDHGMEQGSCKKYVYLNKYLGDVNNVKVVYGPAARLRPTDVPETYYSFN 452

Qy        420 SEEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW-LAVRSKSNTNC 478
Db        453 YEALAKNLSCREPNQHFRPYLKPFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSERKYC 512

Qy        479 GGGNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGS 538
Db        513 GSGFHGSDNLFSNMQALFIGYGPAFKHGAEVDSFENIEVYNLMCDLLGLIPAPNNGSHGS 572

Qy        539 LNHLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQE 598
Db        573 LNHLLKKPIYNPSHPKEEGFLSQC----PIKSTSNDLGCTCDPWIVPIKDFEKQLNLTTE 628

Qy        599 EITATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLP 658
Db        629 DVDDIYHMTVPYGRPRILLKQHRVCLLQQQQFLTGYSLDLLMPLWASYTF--LSNDQFSR 686

Qy        659 PTVPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQY-DALITSNLVPM 717
Db        687 DDFSNCLYQDLRIPLSPVHKCSYYKSNSKLSYGFLTPPRLNRVSNHIYSEALLTSNIVPM 746

Qy        718 YEEFRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKH---LANTDV 774
Db        747 YQSFQVIWHYLHDTLLQRYAHERNGINVVSGPVFDFDYDGRYDSLEILKQNSRVIRSQEI 806

Qy        775 PIPTHYFVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTA 834
Db        807 LIPTHFFIVLTSCKQLSETPLECSA-LESSAYILPHRPDNIESCTHGKRESSWVEELLTL 865

Qy        835 HIARVRDVELLTGLDFYQDKVQPVSEILQLKTYLPTF 871
Db        866 HRARVTDVELITGLSFYQDRQESVSELLRLKTHLPIF 902
```

RESULT 3
Q7ZXN7
ID Q7ZXN7 PRELIMINARY; PRT; 874 AA.
AC Q7ZXN7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

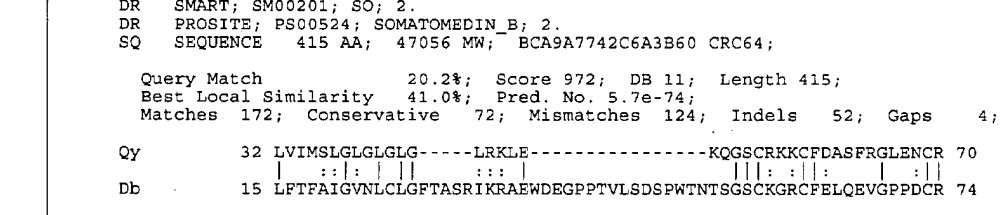
```
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RA   Klein S., Strausberg R.;
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC044675; AAH44675.1; -.
DR   GO; GO:0004519; F:endonuclease activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0009117; P:nucleotide metabolism; IEA.
DR   InterPro; IPR001604; Endonuclease.
DR   InterPro; IPR002591; Phosphodiest.
DR   InterPro; IPR001212; Somatomedin_B.
DR   Pfam; PF01223; Endonuclease; 1.
DR   Pfam; PF01663; Phosphodiest; 1.
DR   Pfam; PF01033; Somatomedin_B; 2.
DR   PRINTS; PR00022; SOMATOMEDINB.
DR   SMART; SM00477; NUC; 1.
DR   SMART; SM00201; SO; 2.
DR   PROSITE; PS00524; SOMATOMEDIN_B; 2.
KW   Hypothetical protein.
SQ   SEQUENCE   874 AA;  99650 MW;  7EA5B4EC1D9FCE46 CRC64;

  Query Match           45.6%;  Score 2191;  DB 13;  Length 874;
  Best Local Similarity   45.1%;  Pred. No. 3.2e-177;
  Matches  409;  Conservative  161;  Mismatches  255;  Indels   82;  Gaps   14;

Qy         15 KNTLKKYKIACIVLLALLVIMSLGLGLGLGLRK----------------LEKQGSCRKKC 58
                ||    :|:  :|   |: : : ||        |                :   |||:::|
Db          4 KNGFSFHKVISLVTFAIGINVCLGFTANRFKRSEEWDEGVASVLSDSPWIRSSGSCKERC 63

Qy         59 FDASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDD 118
               |:         ||||   ||     || ||:: |::: | | | | |||||| | : | ||:|
Db         64 FELIEAEAPACRCDNLCKSYNSCCEDFDEHCLKTGRGWECTKDRCGETRNEENACHCSED 123

Qy        119 CLQKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWD 178
               || | ||| :|: ||:| | | :::|:  :  :|| ||  ||:|:||:||||| |:
Db        124 CLAKGDCCTNYQVVCKGGTHWADDDCEEMKHPECPAGFVRPPLIIFSVDGFRASYMKKGH 183

Qy        179 TLMPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFS 238
              :||||:||::|| || ||| :|||||||| ||: ||||||||||: |:|||   : |||
Db        184 KVMPNIDKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDANFS 243

Qy        239 LSSKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEER 298
              | |:|:  |   || |||:|:||  |||||||:|||   |||:                | |
Db        244 LRSREKFNHRWWGGQPIWITASKQGLKAATFFWP---VAISQ---------------ERR 285

Qy        299 ISTLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQ 358
              | |:|:|| ||   |||  | :| |:|| :||  ||    : : |:| |    | ||:||||
Db        286 IFTVLQWLHLPDNERPYVYALYSEQPDQAGHKYGPFQPELAEQLKVNDKIVGQLMDGLKQ 345

Qy        359 RNLHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHN--IPHDFF 416
               || |||:| : ||||::  | : |::  |  :: | :  |   |:|: |    ||
Db        346 MKLHRCVNVIFVGDHGMEEATCERTEFLNSYLSNVDDFALLPGSIGRMRSRNPANKHD-- 403

Qy        417 SFNSEEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRS---- 472
                 : :| ||:|:||||||||||   |||||||| | ||: :|| ||::|    |
Db        404 ---PKAVVANLTCKKPDQHFKPYLKQHLPKRLHYAFNRRIEDIHLLVDRKWHVARKPMDV 460

Qy        473 -KSNTNCG-GGNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPA 530
               |     |   |:|||:|:  ||: :|| |||||  ||:| |||||:||:|||:| ::||
Db        461 YKRQGKCAFQGDHGYDNKITSMQTVFLGHGPSFKYKTKVPPFENIELYNVMCDVLGLKPA 520

Qy        531 PNNGTHGSLNHLLKVPFYEPSHAEEVSKFSVCGFANPLP--------TESLDCFCPHLQN 582
               |||||||||||:|  |:|:  :|||         |||         | | | |    :
Db        521 SNNGTHGSLNHLLRVASYKPAIPDEVSK--------PLPIVTSPSTVNEELGCSC---DD 569

Qy        583 STQLEQVNQML------NLTQEEITATVK--------VNLPFGRPRVLQKNVDHCLLYHR 628
                  : |::|: |      ::  ||::  :|          || :||| || |   : :|:|
Db        570 KNKAEELNKRLYLKGTDDVAVEELSNEIKELTSRNTDKNLLYGRPAVLYK-TKYSVLHHS 628

Qy        629 EYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPTVPDCLRADVRVPPSESQKCSFYLADKNI 688
              :: ||| ::: ||:|:|||: :  | | :|  : :|:|  | |: |  || || | ||| :
Db        629 DFESGFSESLMMPLWTSYTISKQADVSGIPEHLSNCVRLDPRISPGNSQSCSAYKADKQM 688

Qy        689 THGFLYPPASNRTSDSQYDALITSNLVPMYEEFRKMWDYFHSVLLIKHATERNGVNVVSG 748
              ::|||:||  : ::||:||| : :|::|:|  |:|:|:|  ||: : |||||||||:||
Db        689 SYGFLFPPQLSSSADSKYDAFLITNVIPIYPAFKKIWNYFQRVLVKRFATERNGVNVISG 748

Qy        749 PIFDYNYDGHFDAPDEITKHLANTDVPIPTHYFVVLTSCKNKSHTPENCPGWLDVLPFII 808
              |||||:||| :|  |:| |   :  :|:|||||: ::||| : :   :|| | | |: |:|
Db        749 PIFDYDYDGVYDTMDKI-KMFVDGSIPVPTHYYYIITSCMDFNQAVDNCDGCLSVVSFVI 807

Qy        809 PHRPTNVESCPEGKPEALWVEERFTAHIARVRDVELLTGLDFYQDKVQPVSEILQLKTYL 868
              |||| | |||   : |: |||:    |  |:||:|||||||||:   :  :||| |||||
Db        808 PHRPDNDESCNSSEEESKWVEDLLKMHTTRIRDIELLTGLDFYRKTNRSYTEILSLKTYL 867

Qy        869 PTFETTI 875
                |:|: |
Db        868 HTYESEI 874


RESULT 4
Q9NQM9
ID   Q9NQM9      PRELIMINARY;      PRT;   251 AA.
AC   Q9NQM9;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   DJ1005H11.3 (Phosphodiesterase I/nucleotide pyrophosphatase 3)
DE   (Fragment).
GN   PDNP3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lovell J.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL135904; CAB99214.1; -.
DR   GO; GO:0004519; F:endonuclease activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0009117; P:nucleotide metabolism; IEA.
DR   InterPro; IPR001604; Endonuclease.
DR   InterPro; IPR002591; Phosphodiest.
DR   Pfam; PF01663; Phosphodiest; 1.
DR   SMART; SM00477; NUC; 1.
FT   NON_TER       1      1
FT   NON_TER     251    251
SQ   SEQUENCE   251 AA;  28265 MW;  C726FD4719D4880C CRC64;

  Query Match           28.3%;  Score 1360;  DB 4;  Length 251;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-107;
  Matches  251;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        472 SKSNTNCGGGNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAP 531
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 SKSNTNCGGGNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAP 60

Qy        532 NNGTHGSLNHLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQ 591
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 NNGTHGSLNHLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQ 120

Qy        592 MLNLTQEEITATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQL 651
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 MLNLTQEEITATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQL 180
```

```
Qy        652 GDTSPLPPTVPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALIT 711
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GDTSPLPPTVPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALIT 240

Qy        712 SNLVPMYEEFR 722
              ||||||||||
Db        241 SNLVPMYEEFR 251


RESULT 5
Q9J5H1
ID   Q9J5H1      PRELIMINARY;      PRT;   817 AA.
AC   Q9J5H1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   ORF FPV030 alkaline phosphodiesterase.
GN   FPV030.
OS   Fowlpox virus (FPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Avipoxvirus.
OX   NCBI_TaxID=10261;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20193820; PubMed=10729156;
RA   Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RT   "The genome of fowlpox virus.";
RL   J. Virol. 74:3815-3831(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Afonso C.L., Tulman E.R., Lu Z., Zsak L., Kutish G.F., Rock D.L.;
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF198100; AAF44374.1; -.
DR   GO; GO:0004519; F:endonuclease activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0009117; P:nucleotide metabolism; IEA.
DR   InterPro; IPR001604; Endonuclease.
DR   InterPro; IPR002591; Phosphodiest.
DR   Pfam; PF01663; Phosphodiest; 1.
DR   SMART; SM00477; NUC; 1.
SQ   SEQUENCE   817 AA;  94038 MW;  8404FD00641DA022 CRC64;

  Query Match             26.3%;  Score 1263;  DB 12;  Length 817;
  Best Local Similarity   36.9%;  Pred. No. 2.6e-98;
  Matches  271;  Conservative  145;  Mismatches  279;  Indels   40;  Gaps   18;

Qy        149 QSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKYMRAMYPTKTFPN 208
              |  ||  |: ||:|| :|:||| :||  |:  :| |  |   |: :  || :||| ||||
Db        108 QVTCPPEFERPPLILIAMNGFRYDYLNKWEKYIPTIKDLMEHGVTAP-MRPVYPTNTFPN 166

Qy        209 HYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSKEQNNPAWWHGQPMWLTAMYQGLKAAT 268
              |:|||||| |||| ||   |  :  |:::|:|     |: |:|:| | |  | |:||
Db        167 LYSIVTGLYPISHGITDNEFIDRGTDIEFTIASEETEEVEWFGGEPIWTTIMKNGFKSAT 226

Qy        269 YFWPGSEVAINGSFPSIYMPYNGSVPFEERISTLLKWLDLPKAERPRFYTMYFEEPDSSG 328
              :|||||:  :   |::|  || |||:||||:|:|:|| :   || || :| ||| |||
Db        227 FFWPGSDKVVPRKRPTMYRSYNKSVPYEERINTVLRWLKMDTGYRPYFYALYLEEPGSSG 286

Qy        329 HAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCNKMEYMTD 388
              :  |   || |||: || |  :||:|||   |  | |:||::|||      |:  : |
Db        287 YEYGTDDERVGKALEKVDKAIALLMKGLKDLQLIGCANLILVSDHGMSNVDPKKIVNLKD 346

Qy        389 YFPRINFFYMYEGPAPRIRAHNIPHDFFSFNSEEIVRNLSCRKPDQHFKPYLTPDLPKRL 448
              |    |   :  |  | |:  |: |    |: : |: : ||   || |       |||||
Db        347 YITN-NDVVIKPGATPVIKPQNLNH-IRLFDYDGIISSTSCVMDDQPFIVSYRSRLPKRL 404

Qy        449 HYAKNVRIDKVHLFVDQQWLAVRSKSN-TNCGGGNHGYNNEFRSMEAIFLAHGPSFKEKT 507
              ||    | : : ::::: | :      |  :  || || :| |: | |:|| :||:| :
Db        405 HYGSGFRTEILGVYLEEGWQSTDENGNLKHRSGGFHGSDNSFQDMTAVFLGYGPAFLDDV 464

Qy        508 EVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNHLLKVPFYEPSHAEEVSKFSVCGFANP 567
              |  |:|||:||:||::| | || |||| ||||::|:   |  :|   :   ::
Db        465 RVPIFDNIELYNMMCEILGINPANNNGTVGSLNNILRNSRY--THVSSLDIITI------ 516

Qy        568 LPTESLDC-----FCPHLQNSTQLEQVNQMLNLTQEEITATVK-----VNLPFGRPRVLQ 617
                 || :|        ||:  |  : :::  :   :| ::  :      |||||:| ||
Db        517 ---ES-ECDRHAYVGDHLKGCT-CKNIDRFSSKGNKEDSSRTRSSSYIYNLPFGKPAVLL 571

Qy        618 KNVDHCLLYHREYVSGFGKAMRMPMWSSYTVP-QLGDTSPLPPTVPDCLRADVRVPPSES 676
                 ||:: :  ||: : |  |:|:|:|:::      |:    |   |   |:||
Db        572 NRHHHCIIKNDNYVTAYSKVNRLPLWTSFSIDITYNSTNIYNKT---CYLQDMRV-MYYK 627

Qy        677 QKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEEFRKMWDYFHSVLLIKH 736
              : | :|   |::|:|:|||   | :|  : :|: :| ||||  |:|:|: | | :||::
Db        628 EPCRYYSTQKDVTYGYLYPA---RATD--FQSLLETNTVPMYRNFKKIWEVFMSSILIEY 682

Qy        737 ATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLA-NTDVPIPTHYFVVLTSCKNKSHTPE 795
              : : |||: ||:|| | :|  |:   |:     |  | ||: |||:|| ||:|  :
Db        683 VQKHHVVNVMVGPVFDSNSNGIRDSDWLISMSSGYNNKVYIPSDYFVILTYCKDKDSSLN 742

Qy        796 NCPGWLDVLPFIIPHRPTNV-ESCP-EGKPEALWVEERFTAHIARVRDVELLTGLDFYQD 853
              :|   :    |::|:  |   |||  |      :| : |: |  |::|:| :| : ||::
Db        743 DCYSNIKTESFVVPNSDTYYNESCSRENITSVSYVRKIFSLHRVRIKDIETVTSMSFYRN 802

Qy        854 KVQPVSEILQLKTYL 868
              :  | :  ||||:
Db        803 VYKTSSNVAYLKTYM 817


RESULT 6
O90761
ID   O90761      PRELIMINARY;      PRT;   817 AA.
AC   O90761;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Alkaline phosphodiesterase I (EC 3.1.4.1).
GN   PC1.
OS   Fowlpox virus (FPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Avipoxvirus.
OX   NCBI_TaxID=10261;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FP-9;
RX   MEDLINE=98325194; PubMed=9658122;
RA   Laidlaw S.M., Anwar M.A., Thomas W., Green P., Shaw K., Skinner M.A.;
RT   "Fowlpox virus encodes non-essential homologs of cellular alpha-SNAP,
RT   PC-1 and an orphan human homolog of a secreted nematode protein.";
RL   J. Virol. 72:6742-6751(1998).
DR   EMBL; AJ006408; CAA07014.1; -.
DR   GO; GO:0004519; F:endonuclease activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0003676; F:nucleic acid binding; IEA.
DR   GO; GO:0004528; F:phosphodiesterase I activity; IEA.
DR   GO; GO:0009117; P:nucleotide metabolism; IEA.
DR   InterPro; IPR001604; Endonuclease.
DR   InterPro; IPR002591; Phosphodiest.
DR   Pfam; PF01663; Phosphodiest; 1.
DR   SMART; SM00477; NUC; 1.
KW   Hydrolase.
SQ   SEQUENCE   817 AA;  94004 MW;  CB68DA0508CC568E CRC64;

  Query Match             26.3%;  Score 1263;  DB 12;  Length 817;
  Best Local Similarity   36.9%;  Pred. No. 2.6e-98;
  Matches  271;  Conservative  145;  Mismatches  279;  Indels   40;  Gaps   18;

Qy        149 QSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKYMRAMYPTKTFPN 208
              |  ||  |: ||:|| :|:||| :||  |:  :| |  |   |: :  || :||| ||||
Db        108 QVTCPPEFERPPLILIAMNGFRYDYLNKWEKYIPTIKDLMEHGVTAP-MRPVYPTNTFPN 166

Qy        209 HYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSKEQNNPAWWHGQPMWLTAMYQGLKAAT 268
              |:|||||| |||| ||   |  :  |:::|:|     |: |:|:| | |  | |:||
Db        167 LYSIVTGLYPISHGITDNEFIDRGTDIEFTIASEETEEVEWFGGEPIWTTIMKNGFKSAT 226
```



RESULT 7
Q8CAF0
ID Q8CAF0 PRELIMINARY; PRT; 415 AA.
AC Q8CAF0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ectonucleotide pyrophosphatase/phosphodiesterase 2.
GN ENPP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Hypothalamus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK038940; BAC30174.1; -.
DR MGD; MGI:1321390; Enpp2.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0009117; P:nucleotide metabolism; IEA.
DR InterPro; IPR002591; Phosphodiest.
DR InterPro; IPR001212; Somatomedin_B.
DR Pfam; PF01663; Phosphodiest; 1.
DR Pfam; PF01033; Somatomedin_B; 2.
DR PRINTS; PR00022; SOMATOMEDINB.
DR SMART; SM00201; SO; 2.
DR PROSITE; PS00524; SOMATOMEDIN_B; 2.
SQ SEQUENCE 415 AA; 47056 MW; BCA9A7742C6A3B60 CRC64;

Query Match 20.2%; Score 972; DB 11; Length 415;
Best Local Similarity 41.0%; Pred. No. 5.7e-74;
Matches 172; Conservative 72; Mismatches 124; Indels 52; Gaps 4;



RESULT 8
Q91VZ7
ID Q91VZ7 PRELIMINARY; PRT; 195 AA.
AC Q91VZ7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Similar to alkaline phosphodiesterase.
GN ENPP3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC006944; AAH06944.1; -.
DR MGD; MGI:2143702; Enpp3.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0009117; P:nucleotide metabolism; IEA.
DR InterPro; IPR002591; Phosphodiest.
DR InterPro; IPR001212; Somatomedin_B.
DR Pfam; PF01663; Phosphodiest; 1.
DR Pfam; PF01033; Somatomedin_B; 2.
DR SMART; SM00201; SO; 2.
DR PROSITE; PS00524; SOMATOMEDIN_B; 2.
SQ SEQUENCE 195 AA; 21541 MW; D64F2CB37EFE4A65 CRC64;

Query Match 17.4%; Score 837; DB 11; Length 195;
Best Local Similarity 75.9%; Pred. No. 5.9e-63;
Matches 145; Conservative 21; Mismatches 25; Indels 0; Gaps 0;

```
Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              |:| | ||||:|:||::|||||| |:||||||||:|||||||||||| |:||||||||||
Db          1 MDSRLALATEEPIKKDSLKKYKILCVVLLALLVIVSLGLGLGLGLRKPEEQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              :| |||| |||| | ||||||||||||||:||:|| || ||||| ||| :||||:||||
Db         61 SSHRGLEGCRCDSGCTGRGDCCWDFEDTCVKSTQIWTCNLFRCGENRLETALCSCADDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              |:|||||||:|||||: |: | | ::|: ||| |||||||||||||||||| || ||
Db        121 QRKDCCADYKTVCQGESPWVTEACASSQEPQCPPGFDLPPVILFSMDGFRAEYLQTWSTL 180

Qy        181 MPNINKLKTCG 191
              :||||||   |
Db        181 LPNINKLSKSG 191
```

RESULT 9
Q9H515
ID Q9H515 PRELIMINARY; PRT; 152 AA.
AC Q9H515;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DJ334I18.1 (Ectonucleotide pyrophosphatase/phosphodiesterase 3)
DE (Fragment).
GN ENPP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kay M.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL157377; CAC05481.1; -.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001604; Endonuclease.
DR Pfam; PF01223; Endonuclease; 1.
DR SMART; SM00477; NUC; 1.
FT NON_TER 1 1
SQ SEQUENCE 152 AA; 17449 MW; E22AD9E08483E932 CRC64;

Query Match 17.4%; Score 835; DB 4; Length 152;
Best Local Similarity 100.0%; Pred. No. 6.1e-63;
Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        724 MWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHYFVV 783
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHYFVV 60

Qy        784 LTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVRDVE 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 LTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVRDVE 120

Qy        844 LLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||
Db        121 LLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 152
```

RESULT 10
Q9FS13
ID Q9FS13 PRELIMINARY; PRT; 479 AA.
AC Q9FS13;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nucleotide pyrophosphatase-like protein.
GN SPNPLP-1.
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Amaranthaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE FROM N.A.
RA Sonoda M., Ide H., Sato T., Nakagawa H.;
RT "Cloning and deduced amino acid sequence of a nitrate reductase
RT inactivator (NRI) from spinach leaves identifies a nucleotide
RT pyrophosphatase.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AB032413; BAB20681.1; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0009117; P:nucleotide metabolism; IEA.
DR InterPro; IPR002591; Phosphodiest.
DR Pfam; PF01663; Phosphodiest; 1.
SQ SEQUENCE 479 AA; 53122 MW; 14B8C05F4FB6FF94 CRC64;

Query Match 15.7%; Score 755.5; DB 10; Length 479;
Best Local Similarity 41.1%; Pred. No. 1.8e-55;
Matches 164; Conservative 63; Mismatches 141; Indels 31; Gaps 8;

```
Qy        159 PPVILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKY-MRAMYPTKTFPNHYTIVTGLY 217
              | ||| | |||| | :  |  |||::|  |  ::  :  ::|| ||||||:||||||
Db         94 PIVILISSDGFRFGYQFKAPT--PNIHRLINNGTSAELGLIPVFPTLTFPNHYSIVTGLY 151

Qy        218 PESHGIIDNNMYDVNLNKNFSLSSKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEV- 276
              |  ||||:|  |      ||::| |  | || |:|:| |    | ||||||||||||
Db        152 PPHHGIINNYFVDPITGDKFSMASHE---PKWWLGEPLWETVAKHGKKAATYFWPGSEVI 208

Qy        277 AINGSFPSIY-MPYNGSVPFEERISTLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVS 335
              :  || | : ||||:|||||: |:: : |||  | | | |:|||:||  ||  ||
Db        209 KPDWDCPSNYCIHYNGSIPFEERVDTVISYFDLPINEIPVFMTLYFEDPDHQGHQVGPDD 268

Qy        336 ARVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINF 395
              : :|:  :|   | |::||::| :   ||| :: |||| |  |: :: |  |:
Db        269 PEITQAVINIDKMIGRLIKGLEKRGVFEDVNIFMVGDHGMVGTCDKKLIFLDDLEIPSNW 328

Qy        396 FYMYE-----GPAPRIRAHNIPHDFFSFNSEEIVRNLSCRKPD------QHFKPYLTPDL 444
              |      | |            |:  |||  ::          :| | ||  ||
Db        329 VQSYTPLLAISPPP-----------EFSPAEIVAKMNSGLQSGKVVNGKHLKVYLKEDL 376

Qy        445 PKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGGGNHGYNNEFRSMEAIFLAHGPSFK 504
              ||||||: : ||  :  :|: :  :  || |  || |||:||: ||  ||:|||| |
Db        377 PKRLHYSDSYRIPPIIGLLDEGYKVEQKDSNKNECGGAHGYDNEYFSMRTIFIAHGPQFA 436

Qy        505 EKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNHLL 543
              :  :|  |||:::|||:  :| :: |||||:    | :|
Db        437 KGRKVPSFENVQIYNLVTSILDVEGAPNNGSVSFPNSVL 475
```

RESULT 11
Q7Z3P5
ID Q7Z3P5 PRELIMINARY; PRT; 274 AA.
AC Q7Z3P5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein DKFZp686P13218.
GN DKFZP686P13218.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Human rectum tumor;
RA Bloecker H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BX537580; CAD97789.1; -.
KW Hypothetical protein.
SQ SEQUENCE 274 AA; 30960 MW; 5C4AB440D2D759DA CRC64;

Query Match 15.5%; Score 746; DB 4; Length 274;
Best Local Similarity 54.5%; Pred. No. 5.3e-55;
Matches 140; Conservative 37; Mismatches 78; Indels 2; Gaps 2;

```
Qy        322 EEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCN 381
Db          3 EEPDSSGHSYGPVSSEVIKALQRVDGMVGMLMDGLKELNLHRCLNLILISDHGMEQGSCK 62

Qy        382 KMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNSEEIVRNLSCRKPDQHFKPYLT 441
Db         63 KYIYLNKYLGDVKNIKVIYGPAAARLRPSDVPDKYYSFNYEGIARNLSCREPNQHFKPYLK 122

Qy        442 PDLPKRLHYAKNVRIDKVHLFVDQQW-LAVRSKSNTNCGGGNHGYNNEFRSMEAIFLAHG 500
Db        123 HFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSERKYCGSGFHGSDNVFSNMQALFVGYG 182

Qy        501 PSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNHLLKVPFYEPSHAEEVSKFS 560
Db        183 PGFKHGIEADTFENIEVYNLMCDLLNLTPAPNNGTHGSLNHLLKNPVYTPKHPKEVHPLV 242

Qy        561 VCGFANPLPTESLDCFC 577
Db        243 QCPFTRN-PRDNLGCSC 258
```

RESULT 12
Q9SU83
ID Q9SU83 PRELIMINARY; PRT; 496 AA.
AC Q9SU83;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nucleotide pyrophosphatase-like protein (EC 3.6.1.9).
GN T16L4.190 OR AT4G29680.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Rose M., Hempel S., Entian K.-D., Bancroft I., Mewes H.W.,
RA Mayer K.F.X., Lemcke K., Schueller C.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AL079344; CAB45328.1; -.
DR EMBL; AL161575; CAB79726.1; -.
DR PIR; T09931; T09931.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004551; F:nucleotide diphosphatase activity; IEA.
DR GO; GO:0009117; P:nucleotide metabolism; IEA.
DR InterPro; IPR002591; Phosphodiest.
DR Pfam; PF01663; Phosphodiest; 1.
KW Hydrolase.
SQ SEQUENCE 496 AA; 54678 MW; 8DC2B4346121D732 CRC64;

Query Match 15.5%; Score 743.5; DB 10; Length 496;
Best Local Similarity 41.0%; Pred. No. 2e-54;
Matches 159; Conservative 71; Mismatches 141; Indels 17; Gaps 9;

```
Qy        157 DLPPVILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKY-MRAMYPTKTFPNHYTIVTG 215
Db        105 DKPVVLLISSDGFR--FGYQFKTKLPSIHRLIANGTEAETGLIPVFPTLTFPNHYSIVTG 162

Qy        216 LYPESHGIIDNNMYDVNLNKNFSLSSKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSE 275
Db        163 LYPAYHGIINNHFVDPETGNVFTMASHE---PEWWLGEPLWETVVNQGLKAATYFWPGSE 219

Qy        276 VAINGSF---PSIYMPYNGSVPFEERISTLLKWLDLPKAERPRFYTMYFEEPDSSGHAGG 332
Db        220 VH-KGSWNCPQGLCQNYNGSVPFDDRVDTILSYFDLPSNEIPSFMTLYFEDPDHQGHQVG 278

Qy        333 PVSARVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCNKMEYMTDYFPR 392
Db        279 PDDPQITEAVVNIDRLIGRLIDGLEKRGIFEDVTMIMVGDHGMVGTCDKKLVVLDDLAPW 338

Qy        393 INF---FYMYEGPAPRIRAHNIPHDFFSFNSEEIVRNLSCRKPD--QHFKPYLTPDLPKR 447
Db        339 IKIPSSWVQYYTPLLAIQPPS-GHDAADIVA-KINEGLSSGKVENGKYLKVYLKEDLPSR 396

Qy        448 LHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGGGNHGYNNEFRSMEAIFLAHGPSFKEKT 507
Db        397 LHYVDSDRIPPIIGLVDEGFKVEQKKSKAKECGGAHGYDNAFFSMRTIFIGHGPMFSKGR 456

Qy        508 EVEPFENIEVYNLMCDLLRIQPAPNNGT 535
Db        457 KVPSFENVQIYNVISSILGLKAAPNNGS 484
```

RESULT 13
Q42974
ID Q42974 PRELIMINARY; PRT; 479 AA.
AC Q42974;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nucleotide pyrophosphatase precursor (EC 3.6.1.9) (OSE4) (Nucleotide
DE pyrophosphatase homolog).
GN OSE4 OR P0419B01.2.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TAINUNG 67;
RA Hsing Y.C., Tsao C.V., Chow T., Hsieh J., Chen Z.;
RL Submitted (APR-1995) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. IR64;
RA Hsing Y.-I.C., Hsu T.-F., Chen Z.-Y., Chung M.-C., Tsou C.-H.,
RA Hsieh J.-S.;
RT "Characterization of a rice early embryogenesis specific gene OSE4.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0419B01.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
CC -!- CATALYTIC ACTIVITY: A DINUCLEOTIDE + H(2)O = 2 MONONUCLEOTIDE.
DR EMBL; U25430; AAA67067.1; -.
DR EMBL; AF245483; AAF65458.1; -.
DR EMBL; AP003244; BAB56086.1; -.
DR PIR; T03293; T03293.
DR Gramene; Q42974; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004551; F:nucleotide diphosphatase activity; IEA.
DR GO; GO:0009117; P:nucleotide metabolism; IEA.
DR InterPro; IPR002591; Phosphodiest.
DR Pfam; PF01663; Phosphodiest; 1.
KW Hydrolase; Signal.

```
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20    479       NUCLEOTIDE PYROPHOSPHATASE.
SQ   SEQUENCE   479 AA;  52198 MW;  8E209C3DA17B7920 CRC64;

  Query Match             15.4%;  Score 737.5;  DB 10;  Length 479;
  Best Local Similarity   42.0%;  Pred. No. 6.3e-54;
  Matches  163;  Conservative   64;  Mismatches  140;  Indels   21;  Gaps   10;

Qy        159 PPVILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKY-MRAMYPTKTFPNHYTIVTGLY 217
              | ||| | ||||  | :   |  |:|::|   |  :   :  ::|| ||||||:| ||||
Db         90 PVVILISSDGFRFGYQHKAAT--PHIHRLIGNGTSAATGLVPIFPTLTFPNHYSIATGLY 147

Qy        218 PESHGIIDNNMYDVNLNKNFSLSSKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVA 277
              | |||||:|   |      |::|| |   | || |:|:|:||  ||::||||||||||||
Db        148 PSSHGIINNYFPDPISGDYFTMSSHE---PKWWLGEPLWVTAADQGIQAATYFWPGSEVK 204

Qy        278 INGSF--PSIY-MPYNGSVPFEERISTLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPV 334
                ||:  |  |   ||||||||||:  :| : |||  | |:| |:|||:|| ||  ||
Db        205 -KGSWDCPDKYCRHYNGSVPFEERVDAILGYFDLPSDEMPQFLTLYFEDPDHQGHQVGPD 263

Qy        335 SARVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCNKMEYMTDYFPRIN 394
              : :|:  :|    | |: ||::| :   ||:||: ||||  |   |: :: :  | ||
Db        264 DPAITEAVVRIDEMIGRLIAGLEERGVFEDVNVILVGDHGMVGTCDKKLVFLDELAPWIK 323

Qy        395 F---FYMYEGPAPRIRAHNIPHDFFSFNS--EEIVRNLSCRKPD--QHFKPYLTPDLPKR 447
              : :  |   ||   | | |      ::   |   | :  :: : ||  ||| |
Db        324 LEEDWVLSMTPLLAIR----PPDDMSLPDVVAKMNEGLGSGKVENGEYLRMYLKEDLPSR 379

Qy        448 LHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGGGNHGYNNEFRSMEAIFLAHGPSFKEKT 507
              |||| : ||  :    :: :     :|: |  || |||:| | ||  ||:|||| |:
Db        380 LHYADSYRIPPIIGLPEEGYKVEMKRSDKNECGGAHGYDNAFFSMRTIFIAHGPRFEGGR 439

Qy        508 EVEPFENIEVYNLMCDLLRIQPAPNNGT 535
              |  |||:|:||::  :| ::|||||:
Db        440 VVPSFENVEIYNVIASILNLEPAPNNGS 467


RESULT 14
Q84WJ3
ID   Q84WJ3      PRELIMINARY;      PRT;   457 AA.
AC   Q84WJ3;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   At4g29690.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shinn P., Chen H., Cheuk R., Kim C.J., Bowser L., Carninci P.,
RA   Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
RA   Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA   Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA   Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA   Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
RA   Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT   "Arabidopsis ORF clones.";
RL   Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BT003154; AAO24586.1; -.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0009117; P:nucleotide metabolism; IEA.
DR   InterPro; IPR002591; Phosphodiest.
DR   Pfam; PF01663; Phosphodiest; 1.
SQ   SEQUENCE   457 AA;  51233 MW;  3F85F9D8A9131D0D CRC64;

  Query Match             14.9%;  Score 713.5;  DB 10;  Length 457;
  Best Local Similarity   38.5%;  Pred. No. 6.4e-52;
  Matches  166;  Conservative   76;  Mismatches  148;  Indels   41;  Gaps   12;

Qy        148 QQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKY-MRAMYPTKTF 206
              :: | |:  : | |:| | ||||  | :  :|  |||: | : |  :|  :  ::|| ||
Db         44 RRPQPPKKLNKPVVLLISCDGFRFGYQFKTET--PNIDLLISRGTEAKTGLIPVFPTMTF 101

Qy        207 PNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSKEQNNPAWWHGQPMWLTAMYQGLKA 266
              ||||:| ||||| ||||| |    |    : |:       || || |:|:|:||: ||| |
Db        102 PNHYSIATGLYPASHGIIMNKFTDPVSGELFN----RNLNPKWWLGEPLWVTAVNQGLMA 157

Qy        267 ATYFWPGSEVAINGSF--PSIY--MPYNGSVPFEERISTLLKWLDLPKAERPRFYTMYFE 322
              |||||||::|   ||:  |  :   ||| ||| |||: |:| : |||: | | |  :||:
Db        158 ATYFWPGADVH-KGSWNCPKGFCKAPYNVSVPLEERVDTILNYFDLPEREIPDFMALYFD 216

Qy        323 EPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCNK 382
              |||  ||  ||    || :|:  ||    | :: ||::| : : |::||| ||||      |
Db        217 EPDIQGHEYGPDDPRVTEAVSKVDKMIGRIIMGLEKRKVFSDVHVILLGDHGMATNCDKK 276

Qy        383 MEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFN-------------SEEIVRN--- 426
              : |: |    |            ||  |: ::      : |             : |:||
Db        277 VIYIDDLADWIKI------PADWIQDYS---PVLAMNPRWGKDVKNPGQKNAELVRKMNE 327

Qy        427 -LSCRK--PDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGGGNH 483
               ||  |    : : ||  :||:||||: : ||  :   | :  :  ::::       | |
Db        328 ALSSGKVANGEFLQVYLKENLPQRLHYSDSSRIPPIIGMVGEGLMVKQNRTYVQECSGTH 387

Qy        484 GYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNHLL 543
              ||:| | || :||: :||  |:   :|  |||::||| : ::| ::|||||:      ||
Db        388 GYDNMFFSMRSIFVGYGPRFRRGIKVPSFENVQVYNAVAEILGLRPAPNNGSSLFTRSLL 447

Qy        544 KVPFYEPSHAE 554
              :|| | | |
Db        448 -LPFGETSQVE 457


RESULT 15
Q9SU82
ID   Q9SU82      PRELIMINARY;      PRT;   457 AA.
AC   Q9SU82;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Nucleotide pyrophosphatase-like protein (EC 3.6.1.9).
GN   T16L4.200 OR AT4G29690.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bevan M., Rose M., Hempel S., Entian K.-D., Bancroft I., Mewes H.W.,
RA   Mayer K.F.X., Lemcke K., Schueller C.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Rose M., Hempel S., Entian K.-D., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   EU Arabidopsis sequencing project;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL079344; CAB45329.1; -.
DR   EMBL; AL161575; CAB79727.1; -.
DR   PIR; T09932; T09932.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0004551; F:nucleotide diphosphatase activity; IEA.
DR   GO; GO:0009117; P:nucleotide metabolism; IEA.
DR   InterPro; IPR002591; Phosphodiest.
DR   Pfam; PF01663; Phosphodiest; 1.
KW   Hydrolase.
SQ   SEQUENCE   457 AA;  51261 MW;  3B95E6C0AE1B1215 CRC64;
```

Query Match 14.8%; Score 712.5; DB 10; Length 457;
Best Local Similarity 38.5%; Pred. No. 7.8e-52;
Matches 166; Conservative 76; Mismatches 148; Indels 41; Gaps 12;

```
Qy        148 QQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKY-MRAMYPTKTF 206
              :: | |:  : | |:| | ||||  | :  :|  |||: | : |  :|  :  ::|| ||
Db         44 RRPQPPKKLNKPVVLLISCDGFRFGYQFKTET--PNIDLLISRGTEAKTGLIPVFPTMTF 101

Qy        207 PNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSKEQNNPAWWHGQPMWLTAMYQGLKA 266
              ||||:| ||||| ||||| |   |    : |:       || || |:|:|:||: ||| |
Db        102 PNHYSIATGLYPASHGIIMNKFTDPVSGELFN----RNLNPKWWLGEPLWVTAVNQGLMA 157

Qy        267 ATYFWPGSEVAINGSF--PSIY--MPYNGSVPFEERISTLLKWLDLPKAERPRFYTMYFE 322
              |||||||::|   ||:  |  :   ||| ||| |||: |:| : |||: | | |  :||:
Db        158 ATYFWPGADVH-KGSWNCPKGFCKAPYNVSVPLEERVDTILNYFDLPEREIPDFMALYFD 216

Qy        323 EPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCNK 382
              |||  ||  ||   || :|:  ||   | :: ||::| : : |::||| ||||      |
Db        217 EPDIQGHEYGPDDPRVTEAVSKVDKMIGRIIMGLEKRKVFSDVHVILLGDHGMVTNCDKK 276

Qy        383 MEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFN------------SEEIVRN--- 426
              : |: |    |        ||  |: ::      : |             : |:||
Db        277 VIYIDDLADWIKI------PADWIQDYS---PVLAMNPRWGKDVKNPGQKNAELVRKMNE 327

Qy        427 -LSCRK--PDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGGGNH 483
               ||  |    :  : ||  :||:||||: : ||  :   | :  :  ::::      | |
Db        328 ALSSGKVANGEFLQVYLKENLPQRLHYSDSSRIPPIIGMVGEGLMVKQNRTYVQECSGTH 387

Qy        484 GYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNHLL 543
              ||:| | || :||: :|| |:   :|  |||::||| : ::| ::||||||:       ||
Db        388 GYDNMFFSMRSIFVGYGPRFRRGIKVPSFENVQVYNAVAEILGLRPAPNNGSSLFTRSLL 447

Qy        544 KVPFYEPSHAE 554
               :|| | |  |
Db        448 -LPFGETSQVE 457
```

Search completed: July 6, 2004, 13:19:49
Job time : 52 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 6, 2004, 12:47:02 ; Search time 18 Seconds
(without alignments)
2531.188 Million cell updates/sec

Title: US-10-005-480A-743
Perfect score: 4804
Sequence: 1 MESTLTLATEQPVKKNTLKK.....QPVSEILQLKTYLPTFETTI 875

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	4804	100.0	875	1	NPP3_HUMAN	O14638	h ectonucle
2	3982	82.9	875	1	NPP3_RAT	P97675	r ectonucle
3	2498.5	52.0	925	1	NPP1_HUMAN	P22413	h ectonucle
4	2434	50.7	906	1	NPP1_MOUSE	P06802	m ectonucle
5	2405	50.1	906	1	NPP1_RAT	Q924c3	r ectonucle
6	2111	43.9	862	1	NPP2_MOUSE	Q9r1e6	m ectonucle
7	2091.5	43.5	863	1	NPP2_HUMAN	Q13822	h ectonucle
8	2043.5	42.5	885	1	NPP2_RAT	Q64610	r ectonucle
9	434.5	9.0	742	1	YCR6_YEAST	P25353	saccharomyc
10	412.5	8.6	493	1	YEB6_YEAST	P39997	saccharomyc
11	301	6.3	61	1	PPD1_BOVIN	P15396	bos taurus
12	130	2.7	476	1	TNAG_HUMAN	Q9ujw2	homo sapien
13	129	2.7	369	1	PP11_HUMAN	P21128	homo sapien
14	126	2.6	2004	1	YP73_YERPE	Q8zdj2	yersinia pe
15	121	2.5	4543	1	LRP1_CHICK	P98157	gallus gall
16	119	2.5	1378	1	WR52_ARATH	Q9fh83	arabidopsis
17	116.5	2.4	758	1	YA93_SCHPO	Q09782	schizosacch
18	115.5	2.4	704	1	SSP2_BOMMO	P20613	bombyx mori
19	114	2.4	1106	1	STC_DROME	P40798	drosophila
20	114	2.4	4544	1	LRP1_HUMAN	Q07954	homo sapien
21	113.5	2.4	368	1	NUGL_HUMAN	Q9y2c4	homo sapien
22	113.5	2.4	1808	1	TENA_CHICK	P10039	gallus gall
23	113	2.4	474	1	KLF4_MOUSE	Q60793	mus musculu
24	112.5	2.3	322	1	PNU1_SCHPO	Q10480	schizosacch
25	112.5	2.3	854	1	LDLR_CRIGR	P35950	cricetulus
26	112.5	2.3	2352	1	MOKC_SCHPO	Q9uul4	schizosacch
27	112	2.3	879	1	LDLR_RAT	P35952	rattus norv
28	111	2.3	329	1	NUC1_YEAST	P08466	saccharomyc
29	111	2.3	662	1	FAS1_SCHAM	P10675	schistocerc
30	111	2.3	1442	1	DPO3_UREPA	Q9pqb4	ureaplasma
31	110	2.3	1700	1	BAR3_CHITE	Q03376	chironomus
32	110	2.3	1926	1	LPH_RABIT	P09849	oryctolagus
33	109.5	2.3	871	1	SUL1_HUMAN	Q8iwu6	homo sapien
34	109	2.3	860	1	LDLR_HUMAN	P01130	homo sapien
35	108.5	2.3	2894	1	YD96_METJA	Q58791	methanococc
36	107.5	2.2	470	1	KLF4_HUMAN	O43474	homo sapien
37	107.5	2.2	1169	1	C9DA_BACTP	O06014	bacillus th
38	107	2.2	887	1	ORP3_HUMAN	Q9h4l5	homo sapien
39	107	2.2	2156	1	RP1_HUMAN	P56715	homo sapien
40	107	2.2	4655	1	LRP2_HUMAN	P98164	homo sapien
41	106.5	2.2	4660	1	LRP2_RAT	P98158	rattus norv
42	106	2.2	1787	1	UVRA_CHLMU	Q9pk60	chlamydia m
43	105	2.2	471	1	MM13_BOVIN	O77656	bos taurus
44	104.5	2.2	1682	1	MSP1_PLAF3	P19598	plasmodium
45	103	2.1	787	1	ITB3_MOUSE	O54890	mus musculu

ALIGNMENTS

RESULT 1
NPP3_HUMAN
ID NPP3_HUMAN STANDARD; PRT; 875 AA.
AC O14638;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ectonucleotide pyrophosphatase/phosphodiesterase 3 (E-NPP 3)
DE (Phosphodiesterase I/nucleotide pyrophosphatase 3) (Phosphodiesterase
DE I beta) (PD-Ibeta) (CD203c antigen) [Includes: Alkaline
DE phosphodiesterase I (EC 3.1.4.1); Nucleotide pyrophosphatase
DE (EC 3.6.1.9) (NPPase)].
GN ENPP3 OR PDNP3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=98008933; PubMed=9344668;
RA Piao J.-H., Goding J.W., Nakamura H., Sano K.;
RT "Molecular cloning and chromosomal localization of PD-Ibeta (PDNP3), a
RT new member of the human phosphodiesterase I genes.";
RL Genomics 45:412-415(1997).
RN [2]
RP SEQUENCE OF 189-875 FROM N.A.
RA Hou S., Wohldmann P., Le T.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
CC -!- FUNCTION: Cleaves a variety of phosphodiester and phosphosulfate
CC bonds including deoxynucleotides, nucleotide sugars, and NAD (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: Hydrolytically removes 5'-nucleotides
CC successively from the 3'-hydroxy termini of 3'-hydroxy-terminated
CC oligo-nucleotides.
CC -!- CATALYTIC ACTIVITY: A dinucleotide + H(2)O = 2 mononucleotide.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -!- SIMILARITY: Contains 2 somatomedin-B type domains.
CC -!- DATABASE: NAME=PROW; NOTE=PROW 1:47-49(2000);
CC WWW="http://www.ncbi.nlm.nih.gov/prow/guide/575821061_g.htm".
CC ---

CC ---
DR EMBL; AF005632; AAC51813.1; -.
DR EMBL; AC005587; AAD05192.1; -.
DR Genew; HGNC:3358; ENPP3.
DR MIM; 602182; -.
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0004551; F:nucleotide diphosphatase activity; TAS.
DR GO; GO:0004528; F:phosphodiesterase I activity; TAS.

```
DR   GO; GO:0006796; P:phosphate metabolism; TAS.
DR   InterPro; IPR001604; Endonuclease.
DR   InterPro; IPR002591; Phosphodiest.
DR   InterPro; IPR001212; Somatomedin_B.
DR   Pfam; PF01223; Endonuclease; 1.
DR   Pfam; PF01663; Phosphodiest; 1.
DR   Pfam; PF01033; Somatomedin_B; 2.
DR   SMART; SM00477; NUC; 1.
DR   SMART; SM00201; SO; 2.
DR   PROSITE; PS00524; SOMATOMEDIN_B; 2.
KW   Glycoprotein; Transmembrane; Repeat; Signal-anchor; Hydrolase;
KW   Multifunctional enzyme; Antigen.
FT   DOMAIN       1     11       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     12    30       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                               (POTENTIAL).
FT   DOMAIN       31    875      EXTRACELLULAR (POTENTIAL).
FT   DOMAIN       50    93       SOMATOMEDIN-B LIKE 1.
FT   DOMAIN       94    138      SOMATOMEDIN-B LIKE 2.
FT   DOMAIN       140   510      PHOSPHODIESTERASE.
FT   DOMAIN       605   875      NUCLEASE.
FT   ACT_SITE     205   205      BY SIMILARITY.
FT   SITE         78    80       CELL ATTACHMENT SITE (POTENTIAL).
FT   CARBOHYD     236   236      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     279   279      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     290   290      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     426   426      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     533   533      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     582   582      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     594   594      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     687   687      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     699   699      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     789   789      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   875 AA;  100096 MW;  329CF41667497BC4 CRC64;

  Query Match             100.0%;  Score 4804;  DB 1;  Length 875;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


RESULT 2
NPP3_RAT
ID   NPP3_RAT       STANDARD;      PRT;   875 AA.
AC   P97675; P70641; P97676; Q63490;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ectonucleotide pyrophosphatase/phosphodiesterase 3 (E-NPP 3)
DE   (Phosphodiesterase I/nucleotide pyrophosphatase 3) (Phosphodiesterase
DE   I beta) (PD-Ibeta) (RB13-6 antigen) (B10) [Includes: Alkaline
DE   phosphodiesterase I (EC 3.1.4.1); Nucleotide pyrophosphatase
DE   (EC 3.6.1.9) (NPPase)].
GN   ENNP3 OR PDNP3.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RX   MEDLINE=95247775; PubMed=7730366;
RA   Deissler H., Lottspeich F., Rajewsky M.F.;
RT   "Affinity purification and cDNA cloning of rat neural differentiation
RT   and tumor cell surface antigen gp130RB13-6 reveals relationship to
RT   human and murine PC-1.";
RL   J. Biol. Chem. 270:9849-9855(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Intestine;
RX   MEDLINE=97250927; PubMed=9096610;
RA   Scott L.J., Delautier D., Meerson N.R., Trugnan G., Goding J.W.,
RA   Maurice M.;
RT   "Biochemical and molecular identification of distinct forms of
RT   alkaline phosphodiesterase I expressed on the apical and basolateral
RT   plasma membrane surfaces of rat hepatocytes.";
RL   Hepatology 25:995-1002(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Small intestine;
RA   Sano K.;
RT   "Molecular cloning of phosphodiesterase I cDNA from rat small
RT   intestine.";
RL   Submitted (MAY-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Cleaves a variety of phosphodiester and phosphosulfate
CC       bonds including deoxynucleotides, nucleotide sugars, and NAD (By
CC       similarity).
CC   -!- CATALYTIC ACTIVITY: Hydrolytically removes 5'-nucleotides
```

```
CC       successively from the 3'-hydroxy termini of 3'-hydroxy-terminated
CC       oligo-nucleotides.
CC   -!- CATALYTIC ACTIVITY: A dinucleotide + H(2)O = 2 mononucleotide.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC   -!- SIMILARITY: Contains 2 somatomedin-B type domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z47987; CAA88029.1; -.
DR   EMBL; U78787; AAB61535.1; -.
DR   EMBL; U78788; AAB61536.1; -.
DR   EMBL; D30649; BAA06333.1; -.
DR   PIR; A57080; A57080.
DR   InterPro; IPR001604; Endonuclease.
DR   InterPro; IPR002591; Phosphodiest.
DR   InterPro; IPR001212; Somatomedin_B.
DR   Pfam; PF01223; Endonuclease; 1.
DR   Pfam; PF01663; Phosphodiest; 1.
DR   Pfam; PF01033; Somatomedin_B; 2.
DR   SMART; SM00477; NUC; 1.
DR   SMART; SM00201; SO; 2.
DR   PROSITE; PS00524; SOMATOMEDIN_B; 2.
KW   Glycoprotein; Transmembrane; Repeat; Signal-anchor; Hydrolase;
KW   Multifunctional enzyme; Antigen; Polymorphism.
FT   DOMAIN        1     11       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     12     30       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                (POTENTIAL).
FT   DOMAIN       31    875       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN       51     94       SOMATOMEDIN-B LIKE 1.
FT   DOMAIN       95    139       SOMATOMEDIN-B LIKE 2.
FT   DOMAIN      141    510       PHOSPHODIESTERASE.
FT   DOMAIN      605    875       NUCLEASE.
FT   ACT_SITE    206    206       BY SIMILARITY.
FT   SITE         79     81       CELL ATTACHMENT SITE (POTENTIAL).
FT   CARBOHYD    237    237       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    280    280       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    289    289       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    533    533       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    574    574       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    594    594       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    702    702       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    789    789       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT     124    124       K -> E.
FT   VARIANT     201    201       M -> V.
FT   VARIANT     596    597       SG -> NR.
FT   CONFLICT    111    111       A -> T (IN REF. 3).
FT   CONFLICT    273    273       P -> L (IN REF. 1).
FT   CONFLICT    475    476       SS -> VP (IN REF. 3).
FT   CONFLICT    814    814       N -> KP (IN REF. 3).
SQ   SEQUENCE   875 AA;  99071 MW;  4205F263E8A933EA CRC64;

  Query Match             82.9%;  Score 3982;  DB 1;  Length 875;
  Best Local Similarity   81.3%;  Pred. No. 1.9e-277;
  Matches  712;  Conservative   77;  Mismatches   85;  Indels    2;  Gaps    2;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLGLRKLEKQ-GSCRKKCF 59
              |:| | ||||:|:||::||:||| | ||||||||:|||||||||||| |:  ||||||||
Db          1 MDSRLALATEEPIKKDSLKRYKILCAVLLALLVIVSLGLGLGLGLGLRKPEEHIGSCRKKCF 60

Qy         60 DASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDC 119
              |:| |||| |||| | ||||||||||||||:||:|| || ||||||||||:||||:|||
Db         61 DSSHRGLEGCRCDSGCTDRGDCCWDFEDTCVKSTQIWTCNSFRCGETRLEAALCSCADDC 120

Qy        120 LQKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDT 179
              ||:|||| |||:||||| |: | | ::|: |||||| ||||||||||||||||| || |
Db        121 LQRKDCCTDYKAVCQGEVPWVTEACASSQEPQCPEGFDQPPVILFSMDGFRAEYLQTWST 180

Qy        180 LMPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSL 239
              |:|||||||||:|||||||||||||||||||||||||||||||||||| |||||||
Db        181 LLPNINKLKTCGLHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVYLNKNFSL 240

Qy        240 SSKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERI 299
              || |::|||| |||:|||||||||||:|:||||:||:|||||:||- |: |||:| ||
Db        241 SSVEKSNPAWWSGQPIWLTAMYQGLKAASYYWPGSDVAVNGSFPNIYRNYSNSVPYESRI 300

Qy        300 STLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQR 359
              :|||:|||||||||| |||:| |||||:|| ||||| ||||||:|| ||||||||||||
Db        301 ATLLQWLDLPKAERPSFYTIYVEEPDSAGHKSGPVSAGVIKALQLVDDAFGMLMEGLKQR 360

Qy        360 NLHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFN 419
              |||||||||:|||||||| |:::|||||||| || ||||:|||||||  ||| |||:||
Db        361 NLHNCVNIIVLADHGMDQTSCDRVEYMTDYFPEIN-FYMYQGPAPRIRTRNIPQDFFTFN 419

Qy        420 SEEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCG 479
              ||||||:|||| |||||||||||||||||||||||||||| ||:|||| |:| ::||
Db        420 SEEIVRDLSCRKSDQHFKPYLTPDLPKRLHYAKNVRIDKVHLMVDRQWLAYRNKGSSNCE 479

Qy        480 GGNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSL 539
              || |||||||:|||||||||||||||| :||||||||||:|||| ||||||||:||||
Db        480 GGTHGYNNEFKSMEAIFLAHGPSFKEKTVIEPFENIEVYNLLCDLLHIQPAPNNGSHGSL 539

Qy        540 NHLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEE 599
              ||||| |||:||||||:|| : ||| ||| :||:| | || | | ||||| |||:  |
Db        540 NHLLKAPFYQPSHAEELSKSAGCGFTTPLPKDSLNCSCLALQTSGQEEQVNQRLNLSGGE 599

Qy        600 ITATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPP 659
              ::|| | |||||||||:||| ||||||||||||||||:|||||||||: |||| |||
Db        600 VSATEKTNLPFGRPRVIQKNKDHCLLYHREYVSGFGKAMKMPMWSSYTVPKPGDTSSLPP 659

Qy        660 TVPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYE 719
              |||||||||||| ||||||||||||:|| |||||||    ::|||||||||||||||:
Db        660 TVPDCLRADVRVDPSESQKCSFYLADQNIDHGFLYPPAIKGNNESQYDALITSNLVPMYK 719

Qy        720 EFRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTH 779
              ||:||||||| |||||:| |||||||||||||||||||||||||||| ::| ||||:|||
Db        720 EFKKMWDYFHKVLLIKYAIERNGVNVVSGPIFDYNYDGHFDAPDEITNYVAGTDVPVPTH 779

Qy        780 YFVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARV 839
              |||||||||||:|||::|||||||||::|||||||||||| | | ||||||| ||||||
Db        780 YFVVLTSCKNKTHTPDSCPGWLDVLPFVVPHRPTNVESCPENKAEDLWVEERFKAHIARV 839

Qy        840 RDVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              |||||||||||||:| |||||||||||||||||| |
Db        840 RDVELLTGLDFYQEKTQPVSEILQLKTYLPTFETII 875


RESULT 3
NPP1_HUMAN
ID   NPP1_HUMAN     STANDARD;      PRT;   925 AA.
AC   P22413; Q9NPZ3; Q9P1P6; Q9UP61; Q9Y6K3;
DT   01-AUG-1991 (Rel. 19, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ectonucleotide pyrophosphatase/phosphodiesterase 1 (E-NPP 1)
DE   (Phosphodiesterase I/nucleotide pyrophosphatase 1) (Plasma-cell
DE   membrane glycoprotein PC-1) [Includes: Alkaline phosphodiesterase I
DE   (EC 3.1.4.1); Nucleotide pyrophosphatase (EC 3.6.1.9) (NPPase)].
GN   ENPP1 OR PDNP1 OR PC1 OR NPPS.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91009202; PubMed=2211644;
RA   Buckley M.F., Loveland K.A., McKinstry W.J., Garson O.M., Goding J.W.;
RT   "Plasma cell membrane glycoprotein PC-1. cDNA cloning of the human
RT   molecule, amino acid sequence, and chromosomal location.";
RL   J. Biol. Chem. 265:17506-17511(1990).
RN   [2]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92246539; PubMed=1315502;
RA   Funakoshi I., Kato H., Horie K., Yano T., Hori Y., Kobayashi H.,
RA   Inoue T., Suzuki H., Fukui S., Tsukahara M., Kajii T.,
RA   Yamashina I.;
RT   "Molecular cloning of cDNAs for human fibroblast nucleotide
RT   pyrophosphatase.";
RL   Arch. Biochem. Biophys. 295:180-187(1992).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Bozzali M., Pizzuti A., Trischitta E.;
RT   "Genomic structure of the human PC1 gene.";
RL   Submitted (APR-1999) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 81-925 FROM N.A.
RA   Peck A.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 145-185 FROM N.A., AND VARIANT GLN-173.
RX   MEDLINE=99408501; PubMed=10480624;
RA   Pizzuti A., Frittitta L., Argiolas A., Baratta R., Goldfine I.D.,
RA   Bozzali M., Ercolino T., Scarlato G., Iacoviello L., Vigneri R.,
RA   Tassi V., Trischitta V.;
RT   "A polymorphism (K121Q) of the human glycoprotein PC-1 gene coding
RT   region is strongly associated with insulin resistance.";
RL   Diabetes 48:1881-1884(1999).
RN   [6]
RP   CHARACTERIZATION.
RX   MEDLINE=95094801; PubMed=8001561;
RA   Belli S.I., Goding J.W.;
RT   "Biochemical characterization of human PC-1, an enzyme possessing
RT   alkaline phosphodiesterase I and nucleotide pyrophosphatase
RT   activities.";
RL   Eur. J. Biochem. 226:433-443(1994).
RN   [7]
RP   ACTIVE SITE.
RX   MEDLINE=95255270; PubMed=7737162;
RA   Belli S.I., Mercuri F.A., Sali A., Goding J.W.;
RT   "Autophosphorylation of PC-1 (alkaline phosphodiesterase I/nucleotide
RT   pyrophosphatase) and analysis of the active site.";
RL   Eur. J. Biochem. 228:669-676(1995).
RN   [8]
RP   VARIANTS OPLL PRO-91 AND PHE-287, AND VARIANTS GLN-173; HIS-268 AND
RP   PRO-779.
RX   MEDLINE=99381541; PubMed=10453738;
RA   Nakamura I., Ikegawa S., Okawa A., Okuda S., Koshizuka Y.,
RA   Kawaguchi H., Nakamura K., Koyama T., Goto S., Toguchida J.,
RA   Matsushita M., Ochi T., Takaoka K., Nakamura Y.;
RT   "Association of the human NPPS gene with ossification of the posterior
RT   longitudinal ligament of the spine (OPLL).";
RL   Hum. Genet. 104:492-497(1999).
CC   -!- FUNCTION: Has a broad specificity and cleaves a variety of
CC       substrates, including phosphodiester bonds of nucleotides and
CC       nucleotide sugars and pyrophosphate bonds of nucleotides and
CC       nucleotide sugars. Can hydrolyze nucleoside 5' triphosphates such
CC       as ATP, GTP, CTP, TTP and UTP to their corresponding
CC       monophosphates with release of pyrophosphate. Can hydrolyze
CC       diadenosine polyphosphates and 3',5'-cAMP to AMP. It may play a
CC       role in the regulation of pyrophosphate production, the regulation
CC       of the availability of nucleotide sugars in the endoplasmic
CC       reticulum and Golgi, and the regulation of purinergic signaling.
CC   -!- CATALYTIC ACTIVITY: Hydrolytically removes 5'-nucleotides
CC       successively from the 3'-hydroxy termini of 3'-hydroxy-terminated
CC       oligo-nucleotides.
CC   -!- CATALYTIC ACTIVITY: A dinucleotide + H(2)O = 2 mononucleotide.
CC   -!- SUBUNIT: Homodimer; disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN PLASMA CELLS AND ALSO IN A NUMBER
CC       OF NON-LYMPHOID TISSUES, INCLUDING THE DISTAL CONVOLUTED TUBULE
CC       OF THE KIDNEY, CHONDROCYTES, AND EPIDIDYMIS.
CC   -!- PTM: Autophosphorylated as part of the catalytic cycle of
CC       phosphodiesterase/pyrophosphatase activity.
CC   -!- PTM: N-glycosylated.
CC   -!- DISEASE: Defects in ENPP1 are a cause of increased susceptibility
CC       for ossification of the posterior longitudinal ligament of the
CC       spine (OPLL), a common form of human myelopathy with a prevalence
CC       of as much as 4% in a variety of ethnic groups.
CC   -!- SIMILARITY: Contains 2 somatomedin-B type domains.
CC   -!- DATABASE: NAME=PROW; NOTE=PROW 2:32-44(2001);
CC       WWW="http://www.ncbi.nlm.nih.gov/prow/guide/31658143_g.htm".
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M57736; AAA63237.1; ALT_INIT.
DR   EMBL; D12485; BAA02054.1; ALT_INIT.
DR   EMBL; AF110304; AAF36094.1; -.
DR   EMBL; AF110280; AAF36094.1; JOINED.
DR   EMBL; AF110281; AAF36094.1; JOINED.
DR   EMBL; AF110283; AAF36094.1; JOINED.
DR   EMBL; AF110284; AAF36094.1; JOINED.
DR   EMBL; AF110285; AAF36094.1; JOINED.
DR   EMBL; AF110286; AAF36094.1; JOINED.
DR   EMBL; AF110287; AAF36094.1; JOINED.
DR   EMBL; AF110288; AAF36094.1; JOINED.
DR   EMBL; AF110289; AAF36094.1; JOINED.
DR   EMBL; AF110290; AAF36094.1; JOINED.
DR   EMBL; AF110291; AAF36094.1; JOINED.
DR   EMBL; AF110292; AAF36094.1; JOINED.
DR   EMBL; AF110293; AAF36094.1; JOINED.
DR   EMBL; AF110294; AAF36094.1; JOINED.
DR   EMBL; AF110295; AAF36094.1; JOINED.
DR   EMBL; AF110296; AAF36094.1; JOINED.
DR   EMBL; AF110297; AAF36094.1; JOINED.
DR   EMBL; AF110298; AAF36094.1; JOINED.
DR   EMBL; AF110299; AAF36094.1; JOINED.
DR   EMBL; AF110300; AAF36094.1; JOINED.
DR   EMBL; AF110301; AAF36094.1; JOINED.
DR   EMBL; AF110302; AAF36094.1; JOINED.
DR   EMBL; AF110303; AAF36094.1; JOINED.
DR   EMBL; AJ242020; CAC39442.1; -.
DR   EMBL; AJ242021; CAC39442.1; JOINED.
DR   EMBL; AJ242022; CAC39442.1; JOINED.
DR   EMBL; AJ242023; CAC39442.1; JOINED.
DR   EMBL; AJ242024; CAC39442.1; JOINED.
DR   EMBL; AJ242025; CAC39442.1; JOINED.
DR   EMBL; AJ242026; CAC39442.1; JOINED.
DR   EMBL; AJ242027; CAC39442.1; JOINED.
DR   EMBL; AJ242028; CAC39442.1; JOINED.
DR   EMBL; AJ242029; CAC39442.1; JOINED.
DR   EMBL; AJ242030; CAC39442.1; JOINED.
DR   EMBL; AJ242031; CAC39442.1; JOINED.
DR   EMBL; AJ242032; CAC39442.1; JOINED.
DR   EMBL; AJ242033; CAC39442.1; JOINED.
DR   EMBL; AJ242034; CAC39442.1; JOINED.
DR   EMBL; AJ242035; CAC39442.1; JOINED.
DR   EMBL; AJ242036; CAC39442.1; JOINED.
DR   EMBL; AJ242037; CAC39442.1; JOINED.
DR   EMBL; AJ242038; CAC39442.1; JOINED.
DR   EMBL; AJ242039; CAC39442.1; JOINED.
DR   EMBL; AJ242040; CAC39442.1; JOINED.
DR   EMBL; AJ242041; CAC39442.1; JOINED.
DR   EMBL; AJ242042; CAC39442.1; JOINED.
DR   EMBL; AJ242043; CAC39442.1; JOINED.
DR   EMBL; AJ242044; CAC39442.1; JOINED.
DR   EMBL; AL117378; CAB99365.1; -.
DR   EMBL; AF067177; AAD38420.1; -.
DR   EMBL; AF067178; AAD38421.1; -.
DR   PIR; A39216; A39216.
DR   Genew; HGNC:3356; ENPP1.
DR   MIM; 173335; -.
```

```
DR   MIM; 602475; -.
DR   GO; GO:0016021; C:integral to membrane; TAS.
DR   GO; GO:0004551; F:nucleotide diphosphatase activity; TAS.
DR   GO; GO:0006091; P:energy pathways; TAS.
DR   GO; GO:0006796; P:phosphate metabolism; TAS.
DR   GO; GO:0007584; P:response to nutrients; TAS.
DR   InterPro; IPR001604; Endonuclease.
DR   InterPro; IPR002591; Phosphodiest.
DR   InterPro; IPR001212; Somatomedin_B.
DR   Pfam; PF01223; Endonuclease; 1.
DR   Pfam; PF01663; Phosphodiest; 1.
DR   Pfam; PF01033; Somatomedin_B; 2.
DR   PRINTS; PR00022; SOMATOMEDINB.
DR   SMART; SM00477; NUC; 1.
DR   SMART; SM00201; SO; 2.
DR   PROSITE; PS00524; SOMATOMEDIN_B; 2.
KW   Glycoprotein; Transmembrane; Repeat; Signal-anchor; Hydrolase;
KW   Multifunctional enzyme; Polymorphism; Disease mutation;
KW   Phosphorylation.
FT   DOMAIN        1     76       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     77     97       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                 (POTENTIAL).
FT   DOMAIN       98    925       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN      104    144       SOMATOMEDIN-B LIKE 1.
FT   DOMAIN      145    188       SOMATOMEDIN-B LIKE 2.
FT   DOMAIN      191    591       PHOSPHODIESTERASE.
FT   DOMAIN      654    925       NUCLEASE.
FT   ACT_SITE    256    256
FT   CARBOHYD    179    179       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    285    285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    341    341       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    477    477       N-LINKED (GLCNAC. . .) (POTENTIAL).

  Query Match             52.0%;  Score 2498.5;  DB 1;  Length 925;
  Best Local Similarity   53.1%;  Pred. No. 4e-171;
  Matches  465;  Conservative  137;  Mismatches  250;  Indels   23;  Gaps   11;

Qy         10 EQPVKK----NTLK---KYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFDAS 62
              |:|::|     | |    ||:  :||   ::   ||  ||     ::  ||: :||: :
Db         57 EEPLEKAARARTAKDPNTYKVLSLVLSVCVLTTILGCIFGLKPSCAKEVKSCKGRCFERT 116

Qy         63 FRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQK 122
              |    ||||| || : |:|| |:::||:|   || |||||||| ||  |||:|||||  |
Db        117 F---GNCRCDAACVELGNCCLDYQETCIEPEHIWTCNKFRCGEKRLTRSLCACSDDCKDK 173

Qy        123 KDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMP 182
              ||| :| |||||| ||:|| |::  : ||| ||: || :|||:||||||||:||  |:|
Db        174 GDCCINYSSVCQGEKSWVEEPCESINEPQCPAGFETPPTLLFSLDGFRAEYLHTWGGLLP 233

Qy        183 NINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSK 242
              |:||| || ::| || :||||||||||:|||||||||||||||| |||  :| :||| ||
Db        234 VISKLKKCGTYTKNMRPVYPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMNASFSLKSK 293

Qy        243 EQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTL 302
              |: || |: |:|:|:|| |||||: |:|||||:| ||| || ||  ||||||||||  :
Db        294 EKFNPEWYKGEPIWVTAKYQGLKSGTFFWPGSDVEINGIFPDIYKMYNGSVPFEERILAV 353

Qy        303 LKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLH 362
              |:|| ||| ||| |||:| |||||||| |||| |||||| ||   ||||:|||: |||
Db        354 LQWLQLPKDERPHFYTLYLEEPDSSGHSYGPVSSEVIKALQRVDGMVGMLMDGLKELNLH 413

Qy        363 NCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNSEE 422
              |:|:||::||||:|  | |  |:  |   :   :  ||| |:|  ::|  ::||| |
Db        414 RCLNLILISDHGMEQGSCKKYIYLNKYLGDVKNIKVIYGPAARLRPSDVPDKYYSFNYEG 473

Qy        423 IVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW-LAVRSKSNTNCGGG 481
              | ||||||:|:|||||||   ||||||:||: ||: :  ::| || ||:        || |
Db        474 IARNLSCREPNQHFKPYLKHFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSERKYCGSG 533

Qy        482 NHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNH 541
              || :| | :|:|:|: :|| ||   | : |||||||||||||| : ||||||||||||
Db        534 FHGSDNVFSNMQALFVGYGPGFKHGIEADTFENIEVYNLMCDLLNLTPAPNNGTHGSLNH 593

Qy        542 LLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQMLNLTQEEI 600
              ||| | | | | :||    | |    | ::| | | | :    :|     |||  |
Db        594 LLKNPVYTPKHPKEVHPLVQCPFTRN-PRDNLGCSCNPSI---LPIEDFQTQFNLTVAEE 649

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
                ||:|||||||    |||   :::||: : : ||:|:|||| :  : |
Db        650 KIIKHETLPYGRPRVLQKENTICLLSQHQFMSGYSQDILMPLWTSYTVDR--NDSFSTED 707

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQY-DALITSNLVPMYE 719
              :||  | |:| |    |||||  :  :::||| ||  |: |   | :||:|:|:|||||:
Db        708 FSNCLYQDFRIPLSPVHKCSFYKNNTKVSYGFLSPPQLNKNSSGIYSEALLTTNIVPMYQ 767

Qy        720 EFRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKH---LANTDVPI 776
              |: :| |||  || |:| ||||||||||:||::|||  |: : : :    : | :: |
Db        768 SFQVIWRYFHDTLLRKYAEERNGVNVVSGPVFDFDYDGRCDSLENLRQKRRVIRNQEILI 827

Qy        777 PTHYFVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHI 836
              |||:|:||||||: | || :|   || | ||:|||  | |||  || :: ||||    |
Db        828 PTHFFIVLTSCKDTSQTPLHCEN-LDTLAFILPHRTDNSESCVHGKHDSSWVEELLMLHR 886

Qy        837 ARVRDVELLTGLDFYQDKVQPVSEILQLKTYLPTF 871
              ||: ||| :||| ||| : :|||:||:|||:||||
Db        887 ARITDVEHITGLSFYQQRKEPVSDILKLKTHLPTF 921


RESULT 4
NPP1_MOUSE
ID   NPP1_MOUSE     STANDARD;      PRT;   906 AA.
AC   P06802;
DT   01-JAN-1988 (Rel. 06, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ectonucleotide pyrophosphatase/phosphodiesterase 1 (E-NPP 1)
DE   (Phosphodiesterase I/nucleotide pyrophosphatase 1) (Plasma-cell
DE   membrane glycoprotein PC-1) (Ly-41) [Includes: Alkaline
DE   phosphodiesterase I (EC 3.1.4.1); Nucleotide pyrophosphatase
DE   (EC 3.6.1.9) (NPPase)].
GN   ENPP1 OR PDNP1 OR PC1 OR NPPS.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RX   MEDLINE=87165906; PubMed=3104326;
RA   van Driel I.R., Goding J.W.;
RT   "Plasma cell membrane glycoprotein PC-1. Primary structure deduced
RT   from cDNA clones.";
RL   J. Biol. Chem. 262:4882-4887(1987).
RN   [2]
RP   REVISIONS TO 24; 46-47; 640-641 AND 692, VARIANTS ARG-651 AND SER-680,
RP   AND ALTERNATIVE SPLICING.
RX   MEDLINE=22117315; PubMed=12121276;
RA   Banakh I., Sali A., Dubljevic V., Grobben B., Slegers H., Goding J.W.;
RT   "Structural basis of allotypes of ecto-nucleotide
RT   pyrophosphatase/phosphodiesterase (plasma cell membrane glycoprotein
RT   PC-1) in the mouse and rat, and analysis of allele-specific
RT   xenogeneic antibodies.";
RL   Eur. J. Immunogenet. 29:307-313(2002).
RN   [3]
RP   FUNCTION, AND SEQUENCE FROM N.A.
RX   MEDLINE=91271356; PubMed=1647027;
RA   Rebbe N.F., Tong B.D., Finley E.M., Hickman S.;
RT   "Identification of nucleotide pyrophosphatase/alkaline
RT   phosphodiesterase I activity associated with the mouse plasma cell
RT   differentiation antigen PC-1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:5192-5196(1991).
RN   [4]
RP   SEQUENCE OF 203-219 FROM N.A.
RX   MEDLINE=86094275; PubMed=3001713;
RA   van Driel I.R., Wilks A.F., Pietersz G.A., Goding J.W.;
RT   "Murine plasma cell membrane antigen PC-1: molecular cloning of cDNA
```

```
RT   and analysis of expression.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:8619-8623(1985).
RN   [5]
RP   PARTIAL SEQUENCE.
RX   MEDLINE=85056299; PubMed=3917281;
RA   Stearne P.A., van Driel I.R., Grego B., Simpson R.J., Goding J.W.;
RT   "The murine plasma cell antigen PC-1: purification and partial amino
RT   acid sequence.";
RL   J. Immunol. 134:443-448(1985).
RN   [6]
RP   DISEASE.
RX   MEDLINE=98324779; PubMed=9662402;
RA   Okawa A., Nakamura I., Goto S., Moriya H., Nakamura Y., Ikegawa S.;
RT   "Mutation in Npps in a mouse model of ossification of the posterior
RT   longitudinal ligament of the spine.";
RL   Nat. Genet. 19:271-273(1998).
CC   -!- FUNCTION: Has a broad specificity and cleaves a variety of
CC       substrates, including phosphodiester bonds of nucleotides and
CC       nucleotide sugars and pyrophosphate bonds of nucleotides and
CC       nucleotide sugars. Can hydrolyze nucleoside 5' triphosphates such
CC       as ATP, GTP, CTP, TTP and UTP to their corresponding
CC       monophosphates with release of pyrophosphate. Can hydrolyze
CC       diadenosine polyphosphates and 3',5'-cAMP to AMP. It may play a
CC       role in the regulation of pyrophosphate production, the regulation
CC       of the availability of nucleotide sugars in the endoplasmic
CC       reticulum and Golgi, and the regulation of purinergic signaling
CC       (By similarity).
CC   -!- CATALYTIC ACTIVITY: Hydrolytically removes 5'-nucleotides
CC       successively from the 3'-hydroxy termini of 3'-hydroxy-terminated
CC       oligo-nucleotides.
CC   -!- CATALYTIC ACTIVITY: A dinucleotide + H(2)O = 2 mononucleotide.
CC   -!- SUBUNIT: Homodimer; disulfide-linked.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=2;
CC         IsoId=P06802-1; Sequence=Displayed;
CC       Name=1;
CC         IsoId=P06802-2; Sequence=VSP_006748;
CC   -!- TISSUE SPECIFICITY: SELECTIVELY EXPRESSED ON THE SURFACE OF
CC       ANTIBODY-SECRETING CELLS.
CC   -!- PTM: Autophosphorylated as part of the catalytic cycle of
CC       phosphodiesterase/pyrophosphatase activity (By similarity).
CC   -!- PTM: The N-terminus is blocked.
CC   -!- DISEASE: Defects in ENPP1 are the cause of the tiptoe walking
CC       (ttw) phenotype. Ttw mice exhibit ossification of the spinal
CC       ligaments.
CC   -!- SIMILARITY: Contains 2 somatomedin-B type domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02700; AAA39893.2; -.
DR   EMBL; M12552; AAA39892.1; -.
DR   PIR; A27410; A27410.
DR   MGD; MGI:97370; Enpp1.
DR   GO; GO:0001501; P:skeletal development; IMP.
DR   InterPro; IPR001604; Endonuclease.
DR   InterPro; IPR002591; Phosphodiest.
DR   InterPro; IPR001212; Somatomedin_B.
DR   Pfam; PF01223; Endonuclease; 1.
DR   Pfam; PF01663; Phosphodiest; 1.
DR   Pfam; PF01033; Somatomedin_B; 2.
DR   PRINTS; PR00022; SOMATOMEDINB.
DR   SMART; SM00477; NUC; 1.
DR   SMART; SM00201; SO; 2.
DR   PROSITE; PS00524; SOMATOMEDIN_B; 2.
KW   Glycoprotein; Transmembrane; Repeat; Signal-anchor; Hydrolase;
KW   Multifunctional enzyme; Polymorphism; Phosphorylation;
KW   Alternative splicing.
FT   DOMAIN        1     58       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     59     79       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                 (POTENTIAL).
FT   DOMAIN       80    906       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN       86    126       SOMATOMEDIN-B LIKE 1.
FT   DOMAIN      127    170       SOMATOMEDIN-B LIKE 2.
FT   DOMAIN      173    573       PHOSPHODIESTERASE.
FT   DOMAIN      635    906       NUCLEASE.
FT   ACT_SITE    238    238       BY SIMILARITY.
FT   CARBOHYD    161    161       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    267    267       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    323    323       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    459    459       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    567    567       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    624    624       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    630    630       Missing (in isoform 1).
FT                                 /FTId=VSP_006748.
FT   VARIANT     651    651       H -> R (IN ENPP1-B).
FT   VARIANT     680    680       R -> S (IN ENPP1-B).
SQ   SEQUENCE   906 AA;  103175 MW;  068D45B0ED0F224D CRC64;

  Query Match             50.7%;  Score 2434;  DB 1;  Length 906;
  Best Local Similarity   51.0%;  Pred. No. 1.6e-166;
  Matches  447;  Conservative  152;  Mismatches  250;  Indels   28;  Gaps    9;

Qy         10 EQPVKK----------NTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCF 59
              |:|::|          ||   ||:  :||   ::   ||   ||      ::   ||: :||
Db         39 EEPLEKAERARPAKDPNT---YKVLSLVLSVCVLTTILGCIFGLKPSCAKEVKSCKGRCF 95

Qy         60 DASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDC 119
              : :|    ||||| ||   |:|| ||::|||| | || ||||||| ||   :|||:|||
Db         96 ERTF---SNCRCDAACVSLGNCCLDFQETCVEPTHIWTCNKFRCGEKRLSRFVCSCADDC 152

Qy        120 LQKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDT 179
                 ||| :| |||| : ||:|| |::    :||  |: || :|||:|||||||:||
Db        153 KTHNDCCINYSSVCQDKKSWVEETCESIDTPECPAEFESPPTLLFSLDGFRAEYLHTWGG 212

Qy        180 LMPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSL 239
              |:|  :||| || ::| || |||||||||||:|||||||||||||| |||  :| :|||
Db        213 LLPVISKLKNCGTYTKNMRPMYPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMNASFSL 272

Qy        240 SSKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERI 299
              |||: || |: |||:|:|| :| :|: |||||||:| |:|  | ||  ||||||||||
Db        273 KSKEKFNPLWYKGQPIWVTANHQEVKSGTYFWPGSDVEIDGILPDIYKVYNGSVPFEERI 332

Qy        300 STLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQR 359
                :|:|| ||  ||| |||:| |||||||: ||||: |||||| ||   ||||:|||
Db        333 LAVLEWLQLPSHERPHFYTLYLEEPDSSGHSHGPVSSEVIKALQKVDRLVGMLMDGLKDL 392

Qy        360 NLHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFN 419
              |  |:|:||::||||:|  | |  |:  |   :|   :  ||| |:|  ::|  ::|||
Db        393 GLDKCLNLILISDHGMEQGSCKKYVYLNKYLGDVNNVKVVYGPAARLRPTDVPETYYSFN 452

Qy        420 SEEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW-LAVRSKSNTNC 478
              | : :|||||:|:|||:||| | ||||||:||: ||: :  ::| || ||:       |
Db        453 YEALAKNLSCREPNQHFRPYLKPFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSERKYC 512

Qy        479 GGGNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGS 538
              | | || :| | :|:|:|: :||:||   ||: |||||||||||||| : ||||||:|||
Db        513 GSGFHGSDNLFSNMQALFIGYGPAFKHGAEVDSFENIEVYNLMCDLLGLIPAPNNGSHGS 572

Qy        539 LNHLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQE 598
              |||||| | | ||| :|    | |   |: : | |  |          ::  : |||| |
Db        573 LNHLLKKPIYNPSHPKEEGFLSQC----PIKSTSNDLGCTCDPWIVPIKDFEKQLNLTTE 628

Qy        599 EITATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLP 658
              ::     : :|:||||:| |   |||  :::::|:   : ||:|:|||   | :
Db        629 DVDDIYHMTVPYGRPRILLKQHHVCLLQQQQFLTGYSLDLLMPLWASYTF--LRNDQFSR 686

Qy        659 PTVPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQY-DALITSNLVPM 717
                :||  |:|:| |  |||:| ::  :::||| ||  || |:  | :|||:|||:|||
```

```
Db        687 DDFSNCLYQDLRIPLSPVHKCSYYKSNSKLSYGFLTPPRLNRVSNHIYSEALLTSNIVPM 746

Qy        718 YEEFRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKH---LANTDV 774

Db        747 YQSFQVIWHYLHDTLLQRYAHERNGINVVSGPVFDFDYDGRYDSLEILKQNSRVIRSQEI 806

Qy        775 PIPTHYFVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTA 834

Db        807 LIPTHPFIVLTSCKQLSETPLECSA-LESSAYILPHRPDNIESCTHGKRESSWVEELLTL 865

Qy        835 HIARVRDVELLTGLDFYQDKVQPVSEILQLKTYLPTF 871

Db        866 HRARVTDVELITGLSFYQDRQESVSELLRLKTHLPIF 902


RESULT 5
NPP1_RAT
ID   NPP1_RAT       STANDARD;      PRT;   906 AA.
AC   Q924C3; Q91XQ3; Q920C8;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ectonucleotide pyrophosphatase/phosphodiesterase 1 (E-NPP 1)
DE   (Phosphodiesterase I/nucleotide pyrophosphatase 1) (Plasma-cell
DE   membrane glycoprotein PC-1) [Includes: Alkaline phosphodiesterase I
DE   (EC 3.1.4.1); Nucleotide pyrophosphatase (EC 3.6.1.9) (NPPase)].
GN   ENPP1 OR PDNP1 OR PC1 OR NPPS.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANTS.
RC   STRAIN=Lewis, Louvain, and Wistar;
RX   MEDLINE=22117315; PubMed=12121276;
RA   Banakh I., Sali A., Dubljevic V., Grobben B., Slegers H., Goding J.W.;
RT   "Structural basis of allotypes of ecto-nucleotide
RT   pyrophosphatase/phosphodiesterase (plasma cell membrane glycoprotein
RT   PC-1) in the mouse and rat, and analysis of allele-specific
RT   xenogeneic antibodies.";
RL   Eur. J. Immunogenet. 29:307-313(2002).
CC   -!- FUNCTION: Has a broad specificity and cleaves a variety of
CC       substrates, including phosphodiester bonds of nucleotides and
CC       nucleotide sugars and pyrophosphate bonds of nucleotides and
CC       nucleotide sugars. Can hydrolyze nucleoside 5' triphosphates such
CC       as ATP, GTP, CTP, TTP and UTP to their corresponding
CC       monophosphates with release of pyrophosphate. Can hydrolyze
CC       diadenosine polyphosphates and 3',5'-cAMP to AMP. It may play a
CC       role in the regulation of pyrophosphate production, the regulation
CC       of the availability of nucleotide sugars in the endoplasmic
CC       reticulum and Golgi, and the regulation of purinergic signaling
CC       (By similarity).
CC   -!- CATALYTIC ACTIVITY: Hydrolytically removes 5'-nucleotides
CC       successively from the 3'-hydroxy termini of 3'-hydroxy-terminated
CC       oligo-nucleotides.
CC   -!- CATALYTIC ACTIVITY: A dinucleotide + H(2)O = 2 mononucleotide.
CC   -!- SUBUNIT: Homodimer; disulfide-linked (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=2;
CC         IsoId=Q924C3-1; Sequence=Displayed;
CC       Name=1;
CC         IsoId=Q924C3-2; Sequence=VSP_006749;
CC   -!- PTM: Autophosphorylated as part of the catalytic cycle of
CC       phosphodiesterase/pyrophosphatase activity (By similarity).
CC   -!- SIMILARITY: Contains 2 somatomedin-B type domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF340185; AAK69653.1; -.
DR   EMBL; AF340186; AAK69654.1; -.
DR   EMBL; AF320054; AAL26912.1; -.
DR   InterPro; IPR001604; Endonuclease.
DR   InterPro; IPR002591; Phosphodiest.
DR   InterPro; IPR001212; Somatomedin_B.
DR   Pfam; PF01223; Endonuclease; 1.
DR   Pfam; PF01663; Phosphodiest; 1.
DR   Pfam; PF01033; Somatomedin_B; 2.
DR   PRINTS; PR00022; SOMATOMEDINB.
DR   SMART; SM00477; NUC; 1.
DR   SMART; SM00201; SO; 2.
DR   PROSITE; PS00524; SOMATOMEDIN_B; 2.
KW   Glycoprotein; Transmembrane; Repeat; Signal-anchor; Hydrolase;
KW   Multifunctional enzyme; Polymorphism; Phosphorylation;
KW   Alternative splicing.
FT   DOMAIN        1     58       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     59     79       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                 (POTENTIAL).
FT   DOMAIN       80    906       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN       86    126       SOMATOMEDIN-B LIKE 1.
FT   DOMAIN      127    170       SOMATOMEDIN-B LIKE 2.
FT   DOMAIN      173    573       PHOSPHODIESTERASE.
FT   DOMAIN      635    906       NUCLEASE.
FT   ACT_SITE    238    238       BY SIMILARITY.
FT   CARBOHYD    161    161       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    267    267       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    323    323       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    459    459       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    567    567       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    624    624       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    630    630       Missing (in isoform 1).
FT                                 /FTId=VSP_006749.
FT   VARIANT     440    442       RPT -> NPP (IN STRAIN WISTAR).
FT   VARIANT     457    457       A -> T (IN STRAIN LEW).
FT   VARIANT     555    555       M -> I (IN STRAIN LEW).
FT   VARIANT     568    568       E -> G (IN STRAIN WISTAR).
FT   VARIANT     583    583       T -> I (IN STRAIN LEW).
FT   VARIANT     592    592       F -> V (IN STRAIN LEW).
FT   VARIANT     624    624       N -> K (IN STRAIN LEW).
FT   VARIANT     640    640       N -> H (IN STRAIN LEW).
FT   VARIANT     774    774       V -> I (IN STRAIN LEW).
FT   VARIANT     806    806       N -> I (IN STRAIN LEW).
FT   VARIANT     850    850       T -> I (IN STRAIN LEW).
FT   VARIANT     898    898       H -> Q (IN STRAIN LEW).
SQ   SEQUENCE   906 AA;  102942 MW;  71F56780B279A919 CRC64;

  Query Match           50.1%;  Score 2405;  DB 1;  Length 906;
  Best Local Similarity   50.9%;  Pred. No. 1.9e-164;
  Matches  445;  Conservative  146;  Mismatches  261;  Indels   22;  Gaps    9;

Qy         10 EQPVKK----NTLK---KYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFDAS 62

Db         39 EEPLEKAERARTAKDPNTYKVLSLVLSVCVLTTILGCIFGLKPSCAKEVKSCKGRCFERT 98

Qy         63 FRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQK 122

Db         99 F---SNCRCDAACVSLGNCCLDFQETCVEPTHIWTCNKFRCGEKRLSRFVCSCADDCKAH 155

Qy        123 KDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMP 182

Db        156 NDCCINYSSVCQEKKSWVEEACETIDAPQCPAEFESPPTLLFSLDGFRAEYLHTWGGLLP 215

Qy        183 NINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSK 242

Db        216 VISKLKNCGTYTKNMRPVYPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMNASFSLKSK 275

Qy        243 EQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTL 302

Db        276 EKFNPLWYKGQPIWVTANHQEVRSGTYFWPGSDVEIDGILPDIYKVYNGSVPFEERILAV 335
```

```
Qy        303 LKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLH 362
              |:|| ||  ||| |||:| ||||||||: ||||: |||||| |||  ||||:|||   |
Db        336 LEWLQLPSYERPHFYTLYLEEPDSSGHSHGPVSSEVIKALQKVDHIVGMLMDGLKDLGLD 395

Qy        363 NCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNSEE 422
              |:|:||::||||:|  | |  |:  |   :|   :  ||| |:|   :|  ::||| |
Db        396 KCLNLILISDHGMEQGSCKKYVYLNKYLGDVNNVKVVYGPAARLRPTEVPETYYSFNYEA 455

Qy        423 IVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW-LAVRSKSNTNCGGG 481
              : :|||||: :|||:|||   ||||||:||| ||: :  ::| || ||:          || |
Db        456 LAKNLSCRETNQHFRPYLKHFLPKRLHFAKNDRIEPLTFYLDPQWQLALNPSERKYCGSG 515

Qy        482 NHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNH 541
              || :| | :|:|:|: :||:||   ||: |||||||||||||| : ||||| :||||||
Db        516 FHGSDNLFSNMQALFIGYGPAFKHGAEVDSFENIEVYNLMCDLLGLIPAPNNESHGSLNH 575

Qy        542 LLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEIT 601
              ||| | | ||| :| |  | |    |: : | |  |    :   :    :  ||| : :
Db        576 LLKKPIYTPSHPKEESFLSQC----PIKSVSSDLGCTCDPSIVPIMDFEKQFNLTTDAVE 631

Qy        602 ATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPTV 661
              : :| |||| |||    |||: ::::|:   : ||:|:|||   | :
Db        632 DVYSMTVPNGRPRNLQKQHRVCLLHQQQFLTGYSLDLLMPLWTSYTF--LSNDQFSTDDF 689

Qy        662 PDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQY-DALITSNLVPMYEE 720
              :||  |:|:| |   |||:| :  :::||| ||  || |   | :||:|||:||||:
Db        690 SNCLYQDLRIPLSPMHKCSYYKSTSKLSYGFLTPPRLNRVSRQIYSEALLTSNIVPMYQS 749

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPD---EITKHLANTDVPIP 777
              |: :| | |  :| ::| ||||||||||:||::||| :|: :   : |: : : :  ||
Db        750 FQVIWQYLHDTVLRRYAQERNGVNVVSGPVFDFDYDGRYDSSEILKQNTRVIRSQENLIP 809

Qy        778 THYFVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIA 837
              ||:|:|||||  | :|  |  |:  |::|||| |:||  || |: ||||   | |
Db        810 THFFIVLTSCKQLSESPLKCTA-LESSAFLLPHRPDNIESCTHGKQESAWVEELLALHRA 868

Qy        838 RVRDVELLTGLDFYQDKVQPVSEILQLKTYLPTF 871
              || ||||:||| ||||: : |||:|:|||:|| |
Db        869 RVTDVELITGLSFYQDRQESVSELLRLKTHLPIF 902
```

RESULT 6
NPP2_MOUSE
ID NPP2_MOUSE STANDARD; PRT; 862 AA.
AC Q9R1E6; Q99LG9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ectonucleotide pyrophosphatase/phosphodiesterase 2 (E-NPP 2)
DE (Phosphodiesterase I/nucleotide pyrophosphatase 2) (Phosphodiesterase
DE I alpha) (PD-Ialpha) [Includes: Alkaline phosphodiesterase I
DE (EC 3.1.4.1); Nucleotide pyrophosphatase (EC 3.6.1.9) (NPPase)].
GN ENPP2 OR NPPS2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=20169177; PubMed=10702660;
RA Piao J.-H., Matsuda Y., Nakamura H., Sano K.;
RT "Assignment of Pdnp2, the gene encoding phosphodiesterase I/nucleotide
RT pyrophosphatase 2, to mouse Chromosome 15D2.";
RL Cytogenet. Cell Genet. 87:172-174(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- CATALYTIC ACTIVITY: Hydrolytically removes 5'-nucleotides
CC successively from the 3'-hydroxy termini of 3'-hydroxy-terminated
CC oligo-nucleotides.
CC -!- CATALYTIC ACTIVITY: A dinucleotide + H(2)O = 2 mononucleotide.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -!- SIMILARITY: Contains 2 somatomedin-B type domains.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; AF123542; AAD46480.1; -.
DR EMBL; BC003264; AAH03264.1; -.
DR MGD; MGI:1321390; Enpp2.
DR InterPro; IPR001604; Endonuclease.
DR InterPro; IPR002591; Phosphodiest.
DR InterPro; IPR001212; Somatomedin_B.
DR Pfam; PF01223; Endonuclease; 1.
DR Pfam; PF01663; Phosphodiest; 1.
DR Pfam; PF01033; Somatomedin_B; 2.
DR PRINTS; PR00022; SOMATOMEDINB.
DR SMART; SM00477; NUC; 1.
DR SMART; SM00201; SO; 2.
DR PROSITE; PS00524; SOMATOMEDIN_B; 2.
KW Glycoprotein; Transmembrane; Repeat; Signal-anchor; Hydrolase;
KW Multifunctional enzyme.

FT	DOMAIN	1	11	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	12	30	SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT				(POTENTIAL).
FT	DOMAIN	31	862	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	54	97	SOMATOMEDIN-B LIKE 1.
FT	DOMAIN	98	142	SOMATOMEDIN-B LIKE 2.
FT	DOMAIN	144	501	PHOSPHODIESTERASE.
FT	DOMAIN	597	862	NUCLEASE.
FT	ACT_SITE	209	209	BY SIMILARITY.
FT	SITE	126	128	CELL ATTACHMENT SITE (POTENTIAL).
FT	CARBOHYD	53	53	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	410	410	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	524	524	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	806	806	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	517	517	G -> S (IN REF. 1).
FT	CONFLICT	550	550	P -> T (IN REF. 1).
FT	CONFLICT	573	573	E -> K (IN REF. 1).
FT	CONFLICT	743	743	D -> N (IN REF. 1).

SQ SEQUENCE 862 AA; 98886 MW; C43744DEAA85A355 CRC64;

Query Match 43.9%; Score 2111; DB 1; Length 862;
Best Local Similarity 44.1%; Pred. No. 2.1e-143;
Matches 385; Conservative 171; Mismatches 263; Indels 54; Gaps 10;

```
Qy         32 LVIMSLGLGLGLG-----LRKLE----------------KQGSCRKKCFDASFRGLENCR 70
              |   ::|: | ||     ::: |                  |||: :||:    |  :||
Db         15 LFTFAIGVNLCLGFTASRIKRAEWDEGPPTVLSDSPWTNTSGSCKGRCFELQEVGPPDCR 74
```

```
Qy     71 CDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQKKDCCADYK 130
Db     75 CDNLCKSYSSCCHDFDELCLKTARGWECTKDRCGEVRNEENACHCSEDCLSRGDCCTNYQ 134

Qy    131 SVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMPNINKLKTC 190
Db    135 VVCKGESHWVDDDCEEIRVPECPAGFVRPPLIIFSVDGFRASYMKKGSKVMPNIEKLRSC 194

Qy    191 GIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSKEQNNPAWW 250
Db    195 GTHAPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDATFHLRGREKFNHRWW 254

Qy    251 HGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTLLKWLDLPK 310
Db    255 GGQPLWITATKQGVRAGTFFW-----------------SVSIPHERRILTILQWLSLPD 296

Qy    311 AERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILL 370
Db    297 NERPSVYAFYSEQPDFSGHKYGPFGPEMTNPLREIDKTVGQLMDGLKQLKLHRCVNVIFV 356

Qy    371 ADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNSEEIVRNLSCR 430
Db    357 GDHGMEDVTCDRTEFLSNYLTNVDDITLVPGTLGRIRP-KIPNN-LKYDPKAIIANLTCK 414

Qy    431 KPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW------LAVRSKSNTNC-GGGNH 483
Db    415 KPDQHFKPYMKQHLPKRLHYANNRRIEDLHLLVERRWHVARKPLDVYKKPSGKCFFQGDH 474

Qy    484 GYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNHLL 543
Db    475 GFDNKVNSMQTVFVGYGPTFKYRTKVPPFENIELYNVMCDLLGLKPAPNNGTHGSLNHLL 534

Qy    544 KVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQMLNLTQEEITA 602
Db    535 RTNTFRPTLPEEVSRPNYPGIMYLQSDFDLGCTCDDKVEPKNKLEELNKRLHTK----GS 590

Qy    603 TVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPTVP 662
Db    591 TEERHLLYGRPAVLYR-TSYDILYHTDFESGYSEIFLMPLWTSYTISKQAEVSSIPEHLT 649

Qy    663 DCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEEFR 722
Db    650 NCVRPDVRVSPGFSQNCLAYKNDKQMSYGFLFPPYLSSSPEAKYDAFLVTNMVPMYPAFK 709

Qy    723 KMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHYFV 782
Db    710 RVWTYFQRVLVKKYASERNGVNVISGPIFDYNYDGLRDIEDEIKQYVEGSSIPVPTHYYS 769

Qy    783 VLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVRDV 842
Db    770 IITSCLDFTQPADKCDGPLSVSSFILPHRPDNDESCNSSEDESKWVEELMKMHTARVRDI 829

Qy    843 ELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
Db    830 EHLTGLDFYRKTSRSYSEILTLKTYLHTYESEI 862
```

RESULT 7
NPP2_HUMAN
ID   NPP2_HUMAN   STANDARD;   PRT;   863 AA.
AC   Q13822; Q13827; Q15117;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Ectonucleotide pyrophosphatase/phosphodiesterase 2 (E-NPP 2)
DE   (Phosphodiesterase I/nucleotide pyrophosphatase 2) (Phosphodiesterase
DE   I alpha) (PD-Ialpha) (Autotaxin) [Includes: Alkaline phosphodiesterase
DE   I (EC 3.1.4.1); Nucleotide pyrophosphatase (EC 3.6.1.9) (NPPase)].
GN   ENPP2 OR PDNP2 OR ATX.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 2), PARTIAL SEQUENCE, AND
RP   CHARACTERIZATION.
RC   TISSUE=Melanoma;
RX   MEDLINE=95074054; PubMed=7982964;
RA   Murata J., Lee H.Y., Clair T., Krutzsch H.C., Arestad A.A.,
RA   Sobel M.E., Liotta L.A., Stracke M.L.;
RT   "cDNA cloning of the human tumor motility-stimulating protein,
RT   autotaxin, reveals a homology with phosphodiesterases.";
RL   J. Biol. Chem. 269:30479-30484(1994).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Testis;
RX   MEDLINE=96158950; PubMed=8579579;
RA   Lee H.Y., Murata J., Clair T., Polymeropoulos M.H., Torres R.,
RA   Manrow R.E., Liotta L.A., Stracke M.L.;
RT   "Cloning, chromosomal localization, and tissue expression of autotaxin
RT   from human teratocarcinoma cells.";
RL   Biochem. Biophys. Res. Commun. 218:714-719(1996).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=96163899; PubMed=8586446;
RA   Kawagoe H., Soma O., Goji J., Nishimura N., Narita M., Inazawa J.,
RA   Nakamura H., Sano K.;
RT   "Molecular cloning and chromosomal assignment of the human brain-type
RT   phosphodiesterase I/nucleotide.";
RL   Genomics 30:380-384(1995).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Testis;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: Has a potent tumor cell motility-stimulating activity.
CC   -!- CATALYTIC ACTIVITY: Hydrolytically removes 5'-nucleotides
CC   successively from the 3'-hydroxy termini of 3'-hydroxy-terminated
CC   oligo-nucleotides.
CC   -!- CATALYTIC ACTIVITY: A dinucleotide + H(2)O = 2 mononucleotide.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC   Event=Alternative splicing; Named isoforms=2;
CC   Name=1;
CC   IsoId=Q13822-1; Sequence=Displayed;
CC   Name=2;
CC   IsoId=Q13822-2; Sequence=VSP_006750;
CC   -!- TISSUE SPECIFICITY: Predominantly expressed in brain, placenta,
CC   ovary, and small intestine.
CC   -!- SIMILARITY: Contains 2 somatomedin-B type domains.
CC   --

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; L35594; AAA64785.1; -.
DR   EMBL; L46720; AAB00855.1; -.
DR   EMBL; D45421; BAA08260.1; -.
DR   EMBL; BC034961; AAH34961.1; -.
DR   Genew; HGNC:3357; ENPP2.
DR   MIM; 601060; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0004551; F:nucleotide diphosphatase activity; TAS.
DR   GO; GO:0004528; F:phosphodiesterase I activity; TAS.
DR   GO; GO:0008134; F:transcription factor binding; TAS.
DR   GO; GO:0006928; P:cell motility; TAS.
DR   GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.
DR   GO; GO:0006796; P:phosphate metabolism; TAS.
DR   InterPro; IPR001604; Endonuclease.
DR   InterPro; IPR002591; Phosphodiest.
DR   InterPro; IPR001212; Somatomedin_B.
DR   Pfam; PF01223; Endonuclease; 1.
DR   Pfam; PF01663; Phosphodiest; 1.
DR   Pfam; PF01033; Somatomedin_B; 2.
DR   PRINTS; PR00022; SOMATOMEDINB.
DR   SMART; SM00477; NUC; 1.
DR   SMART; SM00201; SO; 2.
DR   PROSITE; PS00524; SOMATOMEDIN_B; 2.
KW   Glycoprotein; Transmembrane; Repeat; Signal-anchor; Hydrolase;
KW   Multifunctional enzyme; Alternative splicing.
FT   DOMAIN        1     11       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     12     30       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                (POTENTIAL).
FT   DOMAIN       31    863       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN       55     98       SOMATOMEDIN-B LIKE 1.
FT   DOMAIN       99    143       SOMATOMEDIN-B LIKE 2.
FT   DOMAIN      145    502       PHOSPHODIESTERASE.
FT   DOMAIN      598    863       NUCLEASE.
FT   ACT_SITE    210    210       BY SIMILARITY.
FT   SITE        127    129       CELL ATTACHMENT SITE (POTENTIAL).
FT   CARBOHYD     54     54       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    411    411       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    525    525       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    807    807       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    324    324       E -> EESSYGSPPTPAKRPKRKVAPKRRQERPVAPPKKRR
FT                                RKIHRMDHYAAETRQDK (in isoform 2).
FT                                /FTId=VSP_006750.
FT   CONFLICT     23     23       N -> S (IN REF. 1).
FT   CONFLICT     73     73       D -> H (IN REF. 3).
FT   CONFLICT    100    100       G -> A (IN REF. 2).
FT   CONFLICT    291    291       Q -> R (IN REF. 1).
FT   CONFLICT    349    349       H -> R (IN REF. 1).
FT   CONFLICT    629    629       P -> L (IN REF. 1).
FT   CONFLICT    644    644       S -> R (IN REF. 3).
FT   CONFLICT    703    703       V -> A (IN REF. 3).
FT   CONFLICT    769    769       Y -> H (IN REF. 3).
SQ   SEQUENCE   863 AA;  99003 MW;  05B8A2E555C464BC CRC64;

  Query Match             43.5%;  Score 2091.5;  DB 1;  Length 863;
  Best Local Similarity   43.1%;  Pred. No. 5.3e-142;
  Matches  381;  Conservative  176;  Mismatches  267;  Indels   59;  Gaps   10;

Qy         24 ACIVLLALLVIMSLGLGLGLGLRKLEK--------------------QGSCRKKCFDASF 63
              :| ::      : : : ||    ::::                    |||: :||:
Db          9 SCQIISLFTFAVGVNICLGFTAHRIKRAEGWEEGPPTVLSDSPWTNISGSCKGRCPELQE 68

Qy         64 RGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQKK 123
                |  :|||   ||     || ||:: |::: | | | | |||| | | : | ||:||| :
Db         69 AGPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRCGEVRNEENACHCSEDCLARG 128

Qy        124 DCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMPN 183
                ||| :|: ||:||: |::::|:  : ::|| ||  ||:|:|:|||||  |:     :|||
Db        129 DCCTNYQVVCKGESHWVDDDCEEIKAAECPAGFVRPPLIIFSVDGFRASYMKKGSKVMPN 188

Qy        184 INKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSKE 243
                | ||::|| || ||| :|||||||| ||: ||||||||||: |:|||   :  | |  :|
Db        189 IEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDATFHLRGRE 248

Qy        244 QNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTLL 303
              : |  || |||:|:||  ||:|| |:||              |: :|:      | || |:|
Db        249 KFNHRWWGGQPLWITATKQGVKAGTFFW------------SVVIPH------ERRILTIL 290

Qy        304 KWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLHN 363
              :|| ||  |||  |  | |:|| |||  ||     :   |: :|   | ||:|||| ||
Db        291 QWLTLPDHERPSVYAFYSEQPDFSGHKYGPFGPEMTNPLREIDKIVGQLMDGLKQLKLHR 350

Qy        364 CVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIR---AHNIPHDFFSFNS 420
              |||:| : ||||:   |:: |::::|  ::  :  |  |||   ::| :|
Db        351 CVNVIFVGDHGMEDVTCDRTEFLSNYLTNVDDITLVPGTLGRIRSKFSNNAKYD-----P 405

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW------LAVRSKS 474
              : |: ||:|:|||||||||   |||||||| | ||: :|| |:::|       |  |
Db        406 KAIIANLTCKKPDQHFKPYLKQHLPKRLHYANNRRIEDIHLLVERRWHVARKPLDVYKKP 465

Qy        475 NTNC-GGGNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNN 533
              :  |   |:||::|:  ||: :|: :||:|| ||:| ||||||:||:|||| ::|||||
Db        466 SGKCFFQGDHGFDNKVNSMQTVFVGYGPTFKYKTKVPPFENIELYNVMCDLLGLKPAPNN 525

Qy        534 GTHGSLNHLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQM 592
              ||||||||||:   : |:  ||:: :  |          | | |   ::   :|:::|:
Db        526 GTHGSLNHLLRTNTFRPTMPEEVTRPNYPGIMYLQSDFDLGCTCDDKVEPKNKLDELNKR 585

Qy        593 LNLTQEEITATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLG 652
              |:       :| : :| :||| || :   : :||| :: ||: :   ||:|:|||| :
Db        586 LHTK----GSTEERHLLYGRPAVLYR-TRYDILYHTDFESGYSEIFLMPLWTSYTVSKQA 640

Qy        653 DTSPLPPTVPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITS 712
              : | :|  :  |:| |||| || || |  |  || :::|||:||  : : :::||| : :
Db        641 EVSSVPDHLTSCVRPDVRVSPSFSQNCLAYKNDKQMSYGFLFPPYLSSSPEAKYDAFLVT 700

Qy        713 NLVPMYEEFRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANT 772
              |:|||  |:::|:||  ||: |:|:|||||||:|||||||:|||  |  |:|  :::  :
Db        701 NMVPMYPAFKRVWNYFQRVLVKKYASERNGVNVISGPIFDYDYDGLHDTEDKIKQYVEGS 760

Qy        773 DVPIPTHYFVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERF 832
              :|:|||: ::||| : :   : | | | |  ||:|||| | |||   : |: ||||
Db        761 SIPVPTHYYSIITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESKWVEELM 820

Qy        833 TAHIARVRDVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
                | |||||:| || |||::   :   ||| ||||| |:|: |
Db        821 KMHTARVRDIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESEI 863


RESULT 8
NPP2_RAT
ID   NPP2_RAT       STANDARD;      PRT;   885 AA.
AC   Q64610;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ectonucleotide pyrophosphatase/phosphodiesterase 2 (E-NPP 2)
DE   (Phosphodiesterase I/nucleotide pyrophosphatase 2) (Phosphodiesterase
DE   I alpha) (PD-Ialpha) [Includes: Alkaline phosphodiesterase I
DE   (EC 3.1.4.1); Nucleotide pyrophosphatase (EC 3.6.1.9) (NPPase)].
GN   ENPP2 OR NPPS2.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Brain;
RX   MEDLINE=95050605; PubMed=7961762;
RA   Narita M., Goji J., Nakamura H., Sano K.;
RT   "Molecular cloning, expression, and localization of a brain-specific
RT   phosphodiesterase I/nucleotide pyrophosphatase (PD-Ialpha) from rat
RT   brain.";
RL   J. Biol. Chem. 269:28235-28242(1994).
```

```
CC   -!- CATALYTIC ACTIVITY: Hydrolytically removes 5'-nucleotides
CC       successively from the 3'-hydroxy termini of 3'-hydroxy-terminated
CC       oligo-nucleotides.
CC   -!- CATALYTIC ACTIVITY: A dinucleotide + H(2)O = 2 mononucleotide.
CC   -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC   -!- TISSUE SPECIFICITY: Abundantly expressed in cerebrum and
CC       cerebellum. Localized in secretory epithelial cells in the brain
CC       and the eye including choroid plexus epithelial cells, ciliary
CC       epithelial cells, iris pigment epithelial cells, and retinal
CC       pigment cells.
CC   -!- SIMILARITY: Contains 2 somatomedin-B type domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D28560; BAA05910.1; -.
DR   PIR; A55453; A55453.
DR   InterPro; IPR001604; Endonuclease.
DR   InterPro; IPR002591; Phosphodiest.
DR   InterPro; IPR001212; Somatomedin_B.
DR   Pfam; PF01223; Endonuclease; 1.
DR   Pfam; PF01663; Phosphodiest; 1.
DR   Pfam; PF01033; Somatomedin_B; 2.
DR   PRINTS; PR00022; SOMATOMEDINB.
DR   SMART; SM00477; NUC; 1.
DR   SMART; SM00201; SO; 2.
DR   PROSITE; PS00524; SOMATOMEDIN_B; 2.
KW   Glycoprotein; Transmembrane; Repeat; Signal-anchor; Hydrolase;
KW   Multifunctional enzyme.
FT   DOMAIN        1     11       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     12     30       SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT                                (POTENTIAL).
FT   DOMAIN       31    885       EXTRACELLULAR (POTENTIAL).
FT   DOMAIN       54     97       SOMATOMEDIN-B LIKE 1.
FT   DOMAIN       98    142       SOMATOMEDIN-B LIKE 2.
FT   DOMAIN      144    499       PHOSPHODIESTERASE.
FT   DOMAIN      620    885       NUCLEASE.
FT   ACT_SITE    207    207       BY SIMILARITY.
FT   SITE        126    128       CELL ATTACHMENT SITE (POTENTIAL).
FT   CARBOHYD     53     53       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    150    150       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    396    396       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    408    408       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    522    522       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    608    608       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    829    829       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   885 AA;  101309 MW;  8C5C0AFC52007973 CRC64;

  Query Match             42.5%;  Score 2043.5;  DB 1;  Length 885;
  Best Local Similarity   41.7%;  Pred. No. 1.5e-138;
  Matches  380;  Conservative  178;  Mismatches  275;  Indels   79;  Gaps   12;

Qy         14 KKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEK----------------QGSC 54
              ::  |  :::  :   |:    |: : ||    ::::                 |||
Db          3 RQGCLGSFQVISLFTFAI----SVNICLGFTASRIKRAEWDEGPPTVLSDSPWTNTSGSC 58

Qy         55 RKKCFDASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCS 114
              : :||:     |  :|||   ||     || ||:: |::: | | | | || | | : |
Db         59 KGRCFELQEVGPPDCRCDNLCKSYSSCCHDFDELCLKTVRGWECTKDRSGEVRNEENACH 118

Qy        115 CSDDCLQKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDL--PPVILFSMDGFRAE 172
              | :||| : ||| :|: ||:||: |:    | | ::|  |    :  ||:|:||:|||||
Db        119 CPEDCLSRGDCCTNYQVVCKGESHWV----DDAARNQSSECLQVCPPPLIIFSVDGFRAS 174

Qy        173 YLYTWDTLMPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVN 232
              |:     :|||| ||::|| |  | | :||||||| ||: ||||||||||: |:|||
Db        175 YMKKGSKVMPNIEKLRSCGTHVPYTRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPV 234

Qy        233 LNKNFSLSSKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGS 292
              : :| |  :|: |  || |||:|:||  ||::| |:||           : |
Db        235 FDASFHLRGREKFNHRWWGGQPLWITATKQGVRAGTFFW----------------SVS 276

Qy        293 VPFEERISTLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGML 352
              :| | || |:|:|| ||  |||  |  | |:|| ||| ||     :  |: :| | |
Db        277 IPHERRILTILQWLSLPDNERPSVYAFYSEQPDFSGHKYGPFGPEMTNPLREIDKTVGQL 336

Qy        353 MEGLKQRNLHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIP 412
              |:|||  || |||:| : ||||:   |:: |:::|   ::   :  |  |||| :|
Db        337 MDGLKQLRLHRCVNVIFVGDHGMEDVTCDRTEFLSNYLTNVDDITLVPGTLGRIRAKSIN 396

Qy        413 HDFFSFNSEEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW----- 467
              :   :: : |: ||:|:||||||||:   |||||||  | ||: :|| ||::|
Db        397 NS--KYDPKTIIANLTCKKPDQHFKPYMKQHLPKRLHYANNRRIEDIHLLVDRRWHVARK 454

Qy        468 -LAVRSKSNTNC-GGGNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLL 525
               | |  | :  |  |:|||::|:  ||: :|: :||:|| :|:|  |||||:||:|||||
Db        455 PLDVYKKPSGKCFFQGDHGFDNKVNSMQTVFVGYGPTFKYRTKVPPFENIELYNVMCDLL 514

Qy        526 RIQPAPNNGTHGSLNHLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFC-PHLQNST 584
              ::|||||||||||||:   : |:  :|||: : |  |  |  ::
Db        515 GLKPAPNNGTHGSLNHLLRTNTFRPTMPDEVSRPNYPGIMYLQSEFDLGCTCDDKVEPKN 574

Qy        585 QLEQVNQMLNL----------------TQEEITATVKV----NLPFGRPRVLQKNVDHC 623
              :||::|: |:                  :: :  :|:     :| :||| || :  :
Db        575 KLEELNKRLHTKGSTEAETGKFRGSKHENKKNLNGSVEPRKERHLLYGRPAVLYR-TSYD 633

Qy        624 LLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPTVPDCLRADVRVPPSESQKCSFYL 683
              :||| :: ||: :   ||:|:|||: :  : | :|  : :|:|  |||| |  || |  |
Db        634 ILYHTDFESGYSEIFLMPLWTSYTISKQAEVSSIPEHLTNCVRPDVRVSPGFSQNCLAYK 693

Qy        684 ADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEEFRKMWDYFHSVLLIKHATERNGV 743
               || :::|||:||  : : :::||| : :|:|||| |:::| ||  ||: |:|:|||||
Db        694 NDKQMSYGFLFPPYLSSSPEAKYDAFLVTNMVPMYPAFKRVWAYFQRVLVKKYASERNGV 753

Qy        744 NVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHYFVVLTSCKNKSHTPENCPGWLDV 803
              ||:|||||||||||  |  ||| :::  : :|:||||: ::||| : :   : | | | |
Db        754 NVISGPIFDYNYDGLRDTEDEIKQYVEGSSIPVPTHYYSIITSCLDFTQPADKCDGPLSV 813

Qy        804 LPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVRDVELLTGLDFYQDKVQPVSEILQ 863
              ||:|||| | |||   : |: ||||   | ||||||:| |||||||:   :  ||||
Db        814 SSFILPHRPDNDESCNSSEDESKWVEELMKMHTARVRDIEHLTGLDFYRKTSRSYSEILT 873

Qy        864 LKTYLPTFETTI 875
              |||||  |:|: |
Db        874 LKTYLHTYESEI 885


RESULT 9
YCR6_YEAST
ID   YCR6_YEAST     STANDARD;      PRT;   742 AA.
AC   P25353; Q8NIL9;
DT   01-MAY-1992 (Rel. 22, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Hypothetical protein YCR026C.
GN   YCR026C OR YCR26C OR YCR246.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RX   MEDLINE=92244356; PubMed=1574125;
RA   Oliver S.G., van der Aart Q.J.M., Agostoni-Carbone M.L., Aigle M.,
RA   Alberghina L., Alexandraki D., Antoine G., Anwar R., Ballesta J.P.G.,
RA   Benit P., Berben G., Bergantino E., Biteau N., Bolle P.-A.,
RA   Bolotin-Fukuhara M., Brown A.J.P., Brown R., Buhler J.-M.,
RA   Carignani G., Chanet R., Contreras R., Crouzet M., Daignan-Fornier B.,
RA   De Haan M., Defoor E., Delgado M.D., Demolder J., Doira C., Dubois E.,
RA   Dujon B., Duesterhoeft A., Erdmann D., Esteban M., Fabre F.,
```

```
RA   Fairhead C.A., Faye G., Feldmann H., Fiers W.,
RA   Francingues-Gaillard M.-C., Franco L., Frontali L., Fukuhara H.,
RA   Fuller L.J., Gent M.E., Gigot D., Gilliquet V., Glansdorff N.,
RA   Goffeau A., Grenson M., Grisanti P., Grivell L.A., Haasemann M.,
RA   Hatat D., Hegemann J.H., Herbert C.J., Hilger F., Hohmann S.,
RA   Hollenberg C.P., Huse K., Iborra F., Indge K.J., Isono K., Jackman P.,
RA   Jacq C., Jacquet M., James C.M., Jauniaux J.-C., Jia Y., Jimenez A.,
RA   Kleinhans U., Kreisl P., Lafranchi G., Lewis C., van der Linden C.G.,
RA   Lucchini G., Lutzenkirchen K., Maat C., Mannhaupt G., Manzano M.E.,
RA   Martegani E., Mathieu A., Maurer C.T.C., McConnell D., McKee R.A.,
RA   Messenguy F., Mewes H.-W., Molemans F., Montague M.A., Navas L.,
RA   Newlon C.S., Olson M.V., Pallier C., Panzeri L., Pearson B.M.,
RA   Perea J., Philippsen P., Pierard A., Planta R.J., Plevani P.,
RA   Poetsch B., Pohl F.M., Purnelle B., Ramezani Rad M., Rasmussen S.W.,
RA   Raynal A., Remacha M., Richterich P., Roberts A.B., Rodriguez F.,
RA   Sanz E., Schaaff-Gerstenschlaeger I., Scherens B., Schweitzer B.,
RA   Shu Y., Skala J., Slonimski P.P., Sor F., Soustelle C.,
RA   Spiegelberg R., Stateva L.I., Steensma H.Y., Steiner S., Thierry A.,
RA   Thireos G., Triano L.N., Urrestarazu L.A., Valle G., Vetter I.,
RA   van Vliet-Reedijk J.C., Volckaert G., Vreken P., Warmington J.R.,
RA   von Wettstein D., Wicksteed B.L., Wilson C., Wurst H., Xu G.,
RA   Zimmermann F.K., Sgouros J.G.;
RT   "The complete DNA sequence of yeast chromosome III.";
RL   Nature 357:38-46(1992).
RN   [2]
RP   SEQUENCE OF 244-742 FROM N.A.
RX   MEDLINE=92245758; PubMed=1574926;
RA   Bolle P.-A., Gilliquet V., Berben G., Dumont J., Hilger F.;
RT   "The complete sequence of K3B, a 7.9 kb fragment between PGK1 and
RT   CRY1 on chromosome III, reveals the presence of seven open reading
RT   frames.";
RL   Yeast 8:205-213(1992).
RN   [3]
RP   REVISIONS.
RA   Valles G., Volckaerts G.;
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO YEAST YEL016C.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X59720; CAC42978.1; -.
DR   PIR; S19437; S19437.
DR   PIR; S27380; S27380.
DR   GermOnline; 138933; -.
DR   SGD; S0000621; YCR026C.
DR   InterPro; IPR002591; Phosphodiest.
DR   Pfam; PF01663; Phosphodiest; 1.
KW   Hypothetical protein; Transmembrane; Glycoprotein.
FT   TRANSMEM    114    135       POTENTIAL.
FT   CARBOHYD    161    161       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    204    204       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    264    264       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    296    296       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    403    403       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   742 AA;  84733 MW;  83BD5F00D69B09C5 CRC64;

  Query Match             9.0%;  Score 434.5;  DB 1;  Length 742;
  Best Local Similarity   26.5%;  Pred. No. 2.3e-23;
  Matches  122;  Conservative   84;  Mismatches  176;  Indels   79;  Gaps   18;

Qy        159 PPVILFSMDGFRAEYLYTWDT-LMPNINKLKTCG----IHSKYMRAMYPTKTFPNHYTIV 213
              |  |: |:|||   :   :|  : :: :||  |      : :|   :||:|||||:|:|
Db        168 PLTIVISLDGFHPSLISKRNTPFLHDLYELKYDGGMNITSTPFMVPSFPTETFPNHWTLV 227

Qy        214 TGLYPESHGIIDNNMYDVNLNKNFSLSSKEQNNPAWWHG---QPMWLT--AMYQG---LK 265
              || ||  |||: |  :| :||: |     :   |  |:    :|:| |  : : |    |
Db        228 TGQYPIHHGIVSNVFWDPDLNEEFHPGVLD---PRIWNNNDTEPIWQTVQSAFDGDIPFK 284

Qy        266 AATYFWPGSEV----------------AINGSFPSIYMPYNGSVPFEERISTLLKWLDL 308
              |||: ||||:|                |  :  :|   |  :::| :::::|:
Db        285 AATHMWPGSDVNYTKYNEEKLQPEHKNPIARERTPFYFDEFNAKEPLSQKLSKIIEYVDM 344

Qy        309 PKA-ERPRFYTMYFEEPDSSGHAGGPVSAR------VIKALQVVDHAFGMLMEGLKQRNL 361
              |||:    |    |: ||  |  |          : |  ||      |:| |:|||
Db        345 STLNERPQLILGYVPNVDAFGHKHGYPSESEYYYEDFTETLGEVDTFLKQLVESLQERNL 404

Qy        362 HNCVNIILLADHGMDQTYCNK----------MEYMTDYFPRINFFYMYEGPAPRIRAHNI 411
              :  |:::::|||          :   ||   ::  |: |||   |   :
Db        405 TSFTNLVIVSDHGMSDIVVPSNVIIWEDLLDEKLRKDY---VSHAYL-EGPMMAISLKD- 459

Qy        412 PHDFFSFNSEEIVRNLSCRKPDQHFKPYLTPDLPKRLHY--AKNVRIDKVHLFVDQQWLA 469
                | |  |: ||     :  :  |:  : ||  ::   ||  :  :  :  :
Db        460 -----SGNINEVYHNLKTSIDEDKYTVYVNGNFPKEWNFNDGKNHHMASIWIVPEPGYAV 514

Qy        470 V------------RSKSNTNCGG-GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPENI 515
              :            : |:  |     |:|||:|    | ::|:  || | :   :|||:|
Db        515 MKKEQLKKVAKGDHKDKNEDNVFTIGSHGYDNNAIDMRSVFIGMGPYFPQ-GYIEPFQNT 573

Qy        516 EVYNLMCDLLRIQPAPNNGTHGSLNHLLKVPFYEPSHAEEV 556
              |:|||:||:  :      |   |:   :|      ||  :|||
Db        574 EIYNLLCDICGVAEKDRNSNDGT--GMLMNQLREPQSSEEV 612


RESULT 10
YEB6_YEAST
ID   YEB6_YEAST     STANDARD;      PRT;   493 AA.
AC   P39997;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Hypothetical 57.4 kDa protein in PMP2-VAC8 intergenic region.
GN   YEL016C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   MEDLINE=97313264; PubMed=9169868;
RA   Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
RA   Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
RA   Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,
RA   Hunicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
RA   Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,
RA   Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
RA   Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";
RL   Nature 387:78-81(1997).
CC   -!- SIMILARITY: TO YEAST YCR26C.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U18530; AAB64493.1; -.
DR   PIR; S50443; S50443.
DR   GermOnline; 139020; -.
DR   SGD; S0000742; YEL016C.
DR   InterPro; IPR002591; Phosphodiest.
DR   Pfam; PF01663; Phosphodiest; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   493 AA;  57354 MW;  10E67A05C6DEDF09 CRC64;

  Query Match             8.6%;  Score 412.5;  DB 1;  Length 493;
  Best Local Similarity   28.2%;  Pred. No. 5e-22;
```

Matches 126; Conservative 72; Mismatches 154; Indels 95; Gaps 22;

```
Qy        150 SQCPEGFDLPP--------------------VILFSMDGFRAEYL---YTWDTLMPNINK 186
Db         47 SKCPKPIDNGPRTIANRSNTYFNGTHDFKTLTILISIDGFHPRLIDAKYT--PFLYNLHN 104

Qy        187 LKTCG----IHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSK 242
Db        105 LRSPYDMNITTAPYMIPSFPTQTFPNHWSMVTGKYPIEHGIVSNIFWD-----NFTSSEF 159

Qy        243 EQNN---PAWWH-GQPMW--LTAMYQG-LKAATYFWPGSEVAI--NGSFPSIYMP----- 288
Db        160 RPNNLDARIWSNTADPIWQLLQTESQGEYKVATHMWPGSEVVYEDHGDVPRERMPFYFGK 219

Qy        289 YNGSVPFEERISTLLKWLDLPK-AERPRFYTMYFEEPDSSGHAGG--PVSARVIKALQVV 345
Db        220 FNQWEKLQDKLAQIFRYIDMPQLKDRPELVISYIPNVDSYGHSFGYDLRDKRLQKLIGEV 279

Qy        346 DHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCNKMEYMTDY---FPR------INFF 396
Db        280 DGFFLDLIEGLQKRNLLKISNVMIVSDHGMSNVNANDGEHVVVWERVFPADAMSAFISHL 339

Qy        397 YMYEGPAPRIRAHNIPHDFFSFNSEEIVRNLSCRKPDQHFKPYLTPDLPKRLH------- 449
Db        340 Y-NEGPMMMVCLKN-PRD----------KQWICDLIEAQLEKAYGDEISRKFHVILKEDF 387

Qy        450 -------------YAKNVRIDKVHLFVDQQWLAVRSKSNTNCG-GGNHGYN-NEFRSMEA 494
Db        388 DPSWKYFQYDNRKHRYDDRVGDIWILADEYYAIVKEMGDVPIGIMGTHGYNFNNCSDMAS 447

Qy        495 IFLAHGPSFKEKTEVEPFENIEVYNLM 521
Db        448 IFIGMGPMFNNEV-VPPFENIEVYNML 473
```

RESULT 11
PPD1_BOVIN
ID PPD1_BOVIN STANDARD; PRT; 61 AA.
AC P15396;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphodiesterase I (EC 3.1.4.1) (5'-exonuclease) (5'-nucleotide
DE phosphodiesterase) (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=Intestine;
RX MEDLINE=85234541; PubMed=2989287;
RA Culp J.S., Blytt H.J., Hermodson M., Butler L.G.;
RT "Amino acid sequence of the active site peptide of bovine intestinal
RT 5'-nucleotide phosphodiesterase and identification of the active
RT site residue as threonine.";
RL J. Biol. Chem. 260:8320-8324(1985).
CC -!- CATALYTIC ACTIVITY: Hydrolytically removes 5'-nucleotides
CC successively from the 3'-hydroxy termini of 3'-hydroxy-terminated
CC oligo-nucleotides.
DR PIR; A25274; A25274.
DR InterPro; IPR002591; Phosphodiest.
DR Pfam; PF01663; Phosphodiest; 1.
KW Hydrolase.
FT NON_TER 1 1
FT ACT_SITE 39 39 PHOSPHOTHREONINE INTERMEDIATE.
SQ SEQUENCE 61 AA; 6999 MW; 580C8B4807A61C84 CRC64;

Query Match 6.3%; Score 301; DB 1; Length 61;
Best Local Similarity 86.7%; Pred. No. 2.8e-15;
Matches 52; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```
Qy        167 DGFRAEYLYTWDTLMPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDN 226
Db          1 DGFRAEYLQTWSTLVPNINKLKTCGVHSQYLRPAYPTKTFPNHYTIVTGLYPESHGIIDN 60
```

RESULT 12
TNAG_HUMAN
ID TNAG_HUMAN STANDARD; PRT; 476 AA.
AC Q9UJW2; Q9UJW1; Q9ULZ4;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tubulointerstitial nephritis antigen (TIN-Ag).
GN TINAG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE=Kidney;
RX MEDLINE=20119130; PubMed=10652240;
RA Ikeda M., Takemura T., Hino S., Yoshioka K.;
RT "Molecular cloning, expression, and chromosomal localization of a
RT human tubulointerstitial nephritis antigen.";
RL Biochem. Biophys. Res. Commun. 268:225-230(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS TIN1 AND TIN2), TISSUE SPECIFICITY, AND
RP SUBCELLULAR LOCATION.
RC TISSUE=Kidney;
RX MEDLINE=20214190; PubMed=10752525;
RA Zhou B., Nelson T.R., Kashtan C., Gleason B., Michael A.F., Vlassi M.,
RA Charonis A.S.;
RT "Identification of two alternatively spliced forms of human
RT tubulointerstitial nephritis antigen (TIN-Ag).";
RL J. Am. Soc. Nephrol. 11:658-668(2000).
RN [3]
RP IDENTIFICATION, AND DISEASE.
RC TISSUE=Kidney;
RX MEDLINE=87199690; PubMed=3553704;
RA Fliger F.D., Wieslander J., Brentjens J.R., Andres G.A.,
RA Butkowski R.J.;
RT "Identification of a target antigen in human anti-tubular basement
RT membrane nephritis.";
RL Kidney Int. 31:800-887(1987).
RN [4]
RP TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.
RX MEDLINE=92106641; PubMed=1762287;
RA Butkowski R.J., Kleppel M.M., Katz A., Michael A.F., Fish A.J.;
RT "Distribution of tubulointerstitial nephritis antigen and evidence for
RT multiple forms.";
RL Kidney Int. 40:838-846(1991).
RN [5]
RP DISEASE.
RX MEDLINE=93103006; PubMed=1466368;
RA Katz A., Fish A.J., Santamaria P., Nevins T.E., Kim Y.,
RA Butkowski R.J.;
RT "Role of antibodies to tubulointerstitial nephritis antigen in human
RT anti-tubular basement membrane nephritis associated with membranous
RT nephropathy.";
RL Am. J. Med. 93:691-698(1992).
RN [6]
RP FUNCTION.
RX MEDLINE=96366847; PubMed=8770961;
RA Chen Y., Krishnamurti U., Wayner E.A., Michael A.F., Charonis A.S.;
RT "Receptors in proximal tubular epithelial cells for tubulointerstitial
RT nephritis antigen.";
RL Kidney Int. 49:153-157(1996).
RN [7]
RP DEVELOPMENTAL STAGE.
RX MEDLINE=98305799; PubMed=9643646;
RA Nelson T.R., Kim Y., Michael A.F., Butkowski R.J., Charonis A.S.;
RT "Tubulointerstitial nephritis antigen (TIN-ag) is expressed in

```
RT   distinct segments of the developing human nephron.";
RL   Connect. Tissue Res. 37:53-60(1998).
CC   -!- FUNCTION: Mediates adhesion of proximal tubule epithelial cells
CC       via integrins alpha3-beta1 and alphaV-beta3.
CC   -!- SUBCELLULAR LOCATION: Basement membranes.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=TIN1;
CC         IsoId=Q9UJW2-1; Sequence=Displayed;
CC         Note=Major isoform;
CC       Name=TIN2;
CC         IsoId=Q9UJW2-2; Sequence=VSP_050567, VSP_050568;
CC   -!- TISSUE SPECIFICITY: Expressed in the kidney cortex, small
CC       intestine and cornea.
CC   -!- DEVELOPMENTAL STAGE: Initially observed in the Bowman's capsule
CC       during early glomerular capillary loop formation in the kidney. In
CC       more developmentally mature glomeruli, following transition from
CC       early to mid-capillary loop stage, expression is higher in the
CC       proximal tubular basement membrane than in the distal basement
CC       membrane and Bowman's capsule.
CC   -!- DISEASE: Antibodies against TINAG are found in sera of patients
CC       with tubulointerstitial nephritis, a rare autoimmune disorder that
CC       causes acute and chronic renal injury.
CC   -!- SIMILARITY: Belongs to peptidase family C1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB022277; BAA84949.1; -.
DR   EMBL; AF195116; AAF08931.1; -.
DR   EMBL; AF195117; AAF08932.1; -.
DR   PIR; JC7189; JC7189.
DR   HSSP; P07688; 1QDQ.
DR   MEROPS; C01.973; -.
DR   Genew; HGNC:14599; TINAG.
DR   MIM; 606749; -.
DR   GO; GO:0005604; C:basement membrane; IDA.
DR   GO; GO:0005194; F:cell adhesion molecule activity; IDA.
DR   GO; GO:0007155; P:cell adhesion; IDA.
DR   InterPro; IPR000668; Peptidase_C1.
DR   InterPro; IPR009041; PMP_inhibitor.
DR   InterPro; IPR001212; Somatomedin_B.
DR   Pfam; PF00112; Peptidase_C1; 1.
DR   ProDom; PD000158; Peptidase_C1; 1.
DR   SMART; SM00645; Pept_C1; 1.
DR   SMART; SM00201; SO; 1.
KW   Cell adhesion; Extracellular matrix; Basement membrane; Glycoprotein;
KW   Alternative splicing.
FT   SITE        49     50       CLEAVAGE (BY FURIN) (POTENTIAL).
FT   CARBOHYD    38     38       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   175    175       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   314    314       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   360    360       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   455    455       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   119    169       Missing (in isoform TIN2).
FT                               /FTId=VSP_050567.
FT   VARSPLIC   209    300       Missing (in isoform TIN2).
FT                               /FTId=VSP_050568.
FT   CONFLICT     7      8       FS -> IL (IN REF. 1).
FT   CONFLICT   175    175       N -> I (IN REF. 2; AAF08932).
FT   CONFLICT   199    199       P -> L (IN REF. 1).
FT   CONFLICT   333    333       D -> H (IN REF. 1).
FT   CONFLICT   381    381       R -> H (IN REF. 1).
FT   CONFLICT   421    421       L -> R (IN REF. 2; AAF08932).
FT   CONFLICT   437    437       F -> S (IN REF. 1).
FT   CONFLICT   463    463       V -> I (IN REF. 1).
SQ   SEQUENCE   476 AA;  54646 MW;  32DE88E3083C3077 CRC64;
```

```
  Query Match             2.7%;  Score 130;  DB 1;  Length 476;
  Best Local Similarity   21.4%;  Pred. No. 0.083;
  Matches   56;  Conservative   29;  Mismatches   69;  Indels  108;  Gaps   13;

Qy         75 CKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDC-LQKKDCCADYKSVC 133
              |:: | || | :| ||  |  :  |           :|| |   |  :  ||| |||| |
Db         57 CRNFG-CCEDRDDGCV--TEFYAAN----------ALCYCDKFCDRENSDCCPDYKSFC 102

Qy        134 QGETSW----------------------LEENCD----TAQQSQCPEGFDL--PPVI- 162
               : |  |                    ::|||:    : || :| :   |  | :|
Db        103 REEKEWPPHTQPWYPEGCFKDGQHYEEGSVIKENCNSCTCSGQQWKCSQHVCLVRPELIE 162

Qy        163 ------------------LFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKYMRAMYPT 203
                               :  |||:    :   || |:   |       | | |
Db        163 QVNKGDYGWTAQNYSQFWGMTLEDGFK----FRLGTLPPSPMLLS-----MNEMTASLPA 213

Qy        204 KTFPNHYTIVTGLYPE-SHGIID----------------------NNMYDVNLNKNFS 238
              |    : : :  :|  :|| :|                        |  ||:
Db        214 TTDLPEFFVASYKWPGWTHGPLDQKNCAASWAFSTASVAADRIAIQSKGRYTANLSPQNL 273

Qy        239 LSSKEQN---------NPAWWH 251
              :|   :|         : |||:
Db        274 ISCCAKNRHGCNSGSIDRAWWY 295
```

```
RESULT 13
PP11_HUMAN
ID   PP11_HUMAN     STANDARD;      PRT;   369 AA.
AC   P21128;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Placental protein 11 precursor (EC 3.4.21.-) (PP11).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 19-29.
RC   TISSUE=Placenta;
RX   MEDLINE=90274869; PubMed=2350438;
RA   Grundmann U., Roemisch J., Siebold B., Bohn H., Amann E.;
RT   "Cloning and expression of a cDNA encoding human placental protein
RT   11, a putative serine protease with diagnostic significance as a
RT   tumor marker.";
RL   DNA Cell Biol. 9:243-250(1990).
RN   [2]
RP   SOMATOMEDIN-B TYPE DOMAIN.
RX   MEDLINE=91248172; PubMed=1710108;
RA   Jenne D.;
RT   "Homology of placental protein 11 and pea seed albumin 2 with
RT   vitronectin.";
RL   Biochem. Biophys. Res. Commun. 176:1000-1006(1991).
CC   -!- FUNCTION: Probable serine protease.
CC   -!- MISCELLANEOUS: PLACENTAL PROTEIN 11 IS A PLACENTAL-SPECIFIC
CC       PROTEIN BUT IS ALSO ASSOCIATED WITH VARIOUS MALIGNANT NEOPLASMS.
CC   -!- SIMILARITY: Contains 1 somatomedin-B type domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M32402; AAA36464.1; -.
DR   EMBL; M36109; AAA36465.1; -.
DR   PIR; A34614; A34614.
DR   MIM; 606720; -.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0008236; F:serine-type peptidase activity; TAS.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
```

```
DR   InterPro; IPR001212; Somatomedin_B.
DR   Pfam; PF01033; Somatomedin_B; 1.
DR   PRINTS; PR00022; SOMATOMEDINB.
DR   SMART; SM00201; SO; 1.
DR   PROSITE; PS00524; SOMATOMEDIN_B; 1.
KW   Hydrolase; Serine protease; Signal; Polymorphism.
FT   SIGNAL         1     18
FT   CHAIN         19    369       PLACENTAL PROTEIN 11.
FT   DOMAIN        45     89       SOMATOMEDIN-B LIKE.
FT   VARIANT       31     31       E -> Q (in dbSNP:6504).
FT                                 /FTId=VAR_014793.
FT   VARIANT       31     31       E -> V (in dbSNP:6505).
FT                                 /FTId=VAR_014794.
SQ   SEQUENCE   369 AA;  42121 MW;  F5935AE12D7E924C CRC64;

  Query Match             2.7%;  Score 129;  DB 1;  Length 369;
  Best Local Similarity   19.6%;  Pred. No. 0.069;
  Matches   71;  Conservative   48;  Mismatches  104;  Indels  140;  Gaps   17;

Qy         53 SCRKKCFDASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASL 112
               ||: :|::|  |      | |:   |:: |:|| |||  | :
Db         48 SCQGRCYEA-FDKHHQCHCNARCQEFGNCCKDFESLCSDHE------------------ 87

Qy        113 CSCSDDCLQKKDCCADYKSVCQGETSWLE--------ENC----DTAQQ-SQCPEGFDLP 159
                 | | | :  |:: :  :  : : :|:  :       :||     :|  |  :||:
Db         88 VSHSSDAITKEEIQSISEKIYRADTNKAQKEDIVLNSQNCISPSETRNQVDRCPKP---- 143

Qy        160 PVILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPE 219
                 ||:   :  | |::  |     || |                  :|   ||
Db        144 ---LFT---YVNEKLFSKPTYAAFINLL-------------------NNYQRATG---- 173

Qy        220 SHGIIDNNMYDVNLNKNFSLSSKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAIN 279
               ||          ::||     : :        : :  ||: : |
Db        174 -HG------------EHFSAQELAEQDAFL---REIMKTAVMKELY------------- 203

Qy        280 GSFPSIYMPYNGSVPFEERISTLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVI 339
                ||      |      | :  :  :  |    |          |:   || ||||
Db        204 -SFLHHQNRYGSEQEFVDDLKNM--WFGL--------YSRGNEEGDSSG---------- 241

Qy        340 KALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMY 399
                      :|  |     ::    || :                |  : | : ||:  |    |
Db        242 -----FEHVFSGEVKKGKVTGFHNWIRF----------YLEEKEGLVDYYSHI-----Y 280

Qy        400 EGP 402
              :||
Db        281 DGP 283

RESULT 14
YP73_YERPE
ID   YP73_YERPE     STANDARD;      PRT;   2004 AA.
AC   Q8ZDJ2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Hypothetical UPF0192 protein YPO2573/Y1143 precursor.
GN   YPO2573 OR Y1143.
OS   Yersinia pestis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=632;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CO-92 / Biovar Orientalis;
RX   MEDLINE=21470413; PubMed=11586360;
RA   Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA   Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
RA   Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA   Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA   Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA   Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.,
RA   Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT   "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL   Nature 413:523-527(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KIM5 / Biovar Mediaevalis;
RX   MEDLINE=22137863; PubMed=12142430;
RA   Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA   Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA   Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA   Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA   Perry R.D.;
RT   "Genome sequence of Yersinia pestis KIM.";
RL   J. Bacteriol. 184:4601-4611(2002).
CC   -!- SIMILARITY: Belongs to the UPF0192 family.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AJ414152; CAC91375.1; -.
DR   EMBL; AE013717; AAM84721.1; -.
DR   PIR; AC0314; AC0314.
KW   Hypothetical protein; Signal; Complete proteome.
FT   SIGNAL         1     27       POTENTIAL.
FT   CHAIN         28   2004       HYPOTHETICAL PROTEIN YP02573.
SQ   SEQUENCE   2004 AA;  224130 MW;  EA9F8A1F7452672C CRC64;

  Query Match             2.6%;  Score 126;  DB 1;  Length 2004;
  Best Local Similarity   19.7%;  Pred. No. 1.1;
  Matches  131;  Conservative   89;  Mismatches  258;  Indels  186;  Gaps   31;

Qy        236 NFSLSSKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFP-SIYMPYNGSVP 294
              |||:   |:   || |  ||   |  | :|                 ||   ||    |||
Db        283 NFSVVYDEKKLNAWIHSQP--LMPMDKG---------------GSVHLSINKGVNASVA 324

Qy        295 FEERISTLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFG---- 350
                     ||: :      |  |::              ::| ::::    |:  |
Db        325 ATPTTQAQNKWVSV-----PNLYSL----------AVNSINATLVES----DNNNGERAL 365

Qy        351 --MLMEGLKQRNLHNCVNIILLADHGMDQTYCNKMEYMTDYFP---------------- 391
                : : :| : :  | |    || |          |    ||::|
Db        366 IIAISDAVKDKEIKNAVKAWLLPQHNFQAKESAKTS--TDFYPWDMDDIDDNLLQQSTPL 423

Qy        392 --RIN---------FFYMYEGPAPRIRAHNIPHDFFSFNSEEIVRNLSCRKPDQHFKPYL 440
                :|          | : :: |: |     :  :     |    ::       |:: ::
Db        424 ALTLNEAEQEYQPIFSFKFDAPSYRTLLIEVNNSLTSVGGYKM--------PEKIYQIVR 475

Qy        441 TPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGGGNHGYNNEFRSMEAIFLAHG 500
               || || | :          :  :  :  |:| ::| :::  |           : : :    ||
Db        476 VPDYPKTLRFMSQGSL--LSMQGDKQ-ISVAARNMT-------GMKLDIKRVIPSQLQHI 525

Qy        501 PSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNH------------------ 541
               |||                :  : | | ||   | :|:
Db        526 VSFKSSEYSSAHFNRLSDEYFTEHFQYQTALNNDNPGEINYQGVDLSRYLANNPSARRGV 585

Qy        542 -LLKVPFYEP------SHAEE----------VSKFSVCGFANPLPTESLDCFCPHLQNST 584
               || :  ::|          |:||              |:|  |         :  :|  |
Db        586 FLLTLSAWDPEKRDNQQHSEEDYDEDQEWVGDSRFVVITDLGIITKQSQD-----RSRDV 640

Qy        585 QLEQVNQMLNLTQEEITATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWS 644
               :: ::  |         :::       |  :        :|  :        |    :    |  |  |  |:
Db        641 FVQSIHSGLPAADAKVSVVAKNGVV-----LLSQITDSKGHVHFPALDAF-KNERQPV-- 692

Qy        645 SYTVPQLGDTSPLPPTVP-----DCLRADV--RVPPSESQKCSFYL-ADKNITHGFLYPP 696
               : |  : || |  ||         |  |  |:       ||: :   || :|: :
Db        693 MFLVEKEGDVSFLPTRATYDRNLDFSRFDIDGEETPSDPRTLSSYLFSDRGV-------- 744

Qy        697 ASNRTSDSQYDALITSNL--------VPMYEEFRKMWDYFHSVLLIKHATERNGVNVVS- 747
                 |   |     |||               ||:  |  |    |    |  |  |     : :|  | :|
Db        745 --YRPGDRFNIGLITRTANWATALDGVPLRAEIRDPRDTLMSTLPI--TLDSSGFNELSY 800
```

```
Qy        748 -----GPIFDYN-YDGHFDAPDEITKHLANTDVPI----PTHYFVVLTSCKNKSHTPENC 797
                  |  ::| |        :| :  | :| | :.   |    | |         |||
Db        801 TTGENSPTGEWNVYLYLVGKNNETSMLLGHTTVNVKEFEPDRLKVQL------QLTPERQ 854

Qy        798 PGWL 801
                ||:
Db        855 QGWV 858


RESULT 15
LRP1_CHICK
ID   LRP1_CHICK     STANDARD;      PRT;  4543 AA.
AC   P98157;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Low-density lipoprotein receptor-related protein 1 precursor (LRP)
DE   (Alpha-2-macroglobulin receptor) (A2MR).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   STRAIN=White leghorn; TISSUE=Liver, and Ovary;
RX   MEDLINE=94103212; PubMed=7506255;
RA   Nimpf J., Stifani S., Bilous P.T., Schneider W.J.;
RT   "The somatic cell-specific low density lipoprotein receptor-related
RT   protein of the chicken. Close kinship to mammalian low density
RT   lipoprotein receptor gene family members.";
RL   J. Biol. Chem. 269:212-219(1994).
CC   -!- FUNCTION: Involved in the plasma clearance of chylomicron remnants
CC       and activated alpha 2-macroglobulin, as well as the local
CC       metabolism of complexes between plasminogen activators and their
CC       endogenous inhibitors. Binds vitellogenin, calcium and alpha 2-
CC       macroglobulin.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P98157-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P98157-2; Sequence=VSP_004312;
CC   -!- TISSUE SPECIFICITY: Somatic.
CC   -!- PTM: Cleaved into a 85 kDa membrane-spanning subunit (LRP-85) and
CC       a 515 kDa large extracellular domain (LRP-515) that remains non-
CC       covalently associated.
CC   -!- SIMILARITY: Contains 22 EGF-like domains.
CC   -!- SIMILARITY: Contains 31 LDL-receptor class A domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X74904; CAA52870.1; -.
DR   PIR; A53102; A53102.
DR   PDB; 1LPX; 29-DEC-99.
DR   InterPro; IPR000152; Asx_hydroxyl_S.
DR   InterPro; IPR001881; EGF_Ca.
DR   InterPro; IPR006209; EGF_like.
DR   InterPro; IPR002172; LDL_receptor_A.
DR   InterPro; IPR000033; Ldl_receptor_rep.
DR   Pfam; PF00008; EGF; 14.
DR   Pfam; PF00057; ldl_recept_a; 31.
DR   Pfam; PF00058; ldl_recept_b; 33.
DR   PRINTS; PR00261; LDLRECEPTOR.
DR   SMART; SM00179; EGF_CA; 3.
DR   SMART; SM00192; LDLa; 31.
DR   SMART; SM00135; LY; 34.
DR   PROSITE; PS00010; ASX_HYDROXYL; 3.
DR   PROSITE; PS00022; EGF_1; 5.
DR   PROSITE; PS01186; EGF_2; 7.
DR   PROSITE; PS50026; EGF_3; 8.
DR   PROSITE; PS01187; EGF_CA; 2.
DR   PROSITE; PS01209; LDLRA_1; 27.
DR   PROSITE; PS50068; LDLRA_2; 31.
KW   Receptor; Transmembrane; Repeat; Endocytosis; Glycoprotein; Signal;
KW   Calcium-binding; EGF-like domain; Coated pits; Alternative splicing;
KW   3D-structure.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22   4543       LOW-DENSITY LIPOPROTEIN RECEPTOR-RELATED
FT                                PROTEIN 1.
FT   DOMAIN       22   4419       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   4420   4443       POTENTIAL.
FT   DOMAIN     4444   4543       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       27     68       LDL-RECEPTOR CLASS A 1.
FT   DOMAIN       72    112       LDL-RECEPTOR CLASS A 2.
FT   DOMAIN      113    151       EGF-LIKE 1.
FT   DOMAIN      152    191       EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN      476    522       EGF-LIKE 3.
FT   DOMAIN      801    841       EGF-LIKE 4.
FT   DOMAIN      850    890       LDL-RECEPTOR CLASS A 3.
FT   DOMAIN      891    931       LDL-RECEPTOR CLASS A 4.
FT   DOMAIN      932    971       LDL-RECEPTOR CLASS A 5.
FT   DOMAIN      972   1011       LDL-RECEPTOR CLASS A 6.
FT   DOMAIN     1011   1051       LDL-RECEPTOR CLASS A 7.
FT   DOMAIN     1058   1097       LDL-RECEPTOR CLASS A 8.
FT   DOMAIN     1100   1140       LDL-RECEPTOR CLASS A 9.
FT   DOMAIN     1141   1180       LDL-RECEPTOR CLASS A 10.
FT   DOMAIN     1181   1220       EGF-LIKE 5.
FT   DOMAIN     1221   1260       EGF-LIKE 6.
FT   DOMAIN     1534   1577       EGF-LIKE 7.
FT   DOMAIN     1842   1883       EGF-LIKE 8.
FT   DOMAIN     2151   2191       EGF-LIKE 9.
FT   DOMAIN     2472   2512       EGF-LIKE 10.
FT   DOMAIN     2516   2557       LDL-RECEPTOR CLASS A 11.
FT   DOMAIN     2558   2596       LDL-RECEPTOR CLASS A 12.
FT   DOMAIN     2597   2635       LDL-RECEPTOR CLASS A 13.
FT   DOMAIN     2636   2684       LDL-RECEPTOR CLASS A 14.
FT   DOMAIN     2688   2730       LDL-RECEPTOR CLASS A 15.
FT   DOMAIN     2730   2769       LDL-RECEPTOR CLASS A 16.
FT   DOMAIN     2770   2812       LDL-RECEPTOR CLASS A 17.
FT   DOMAIN     2814   2853       LDL-RECEPTOR CLASS A 18.
FT   DOMAIN     2854   2897       LDL-RECEPTOR CLASS A 19.
FT   DOMAIN     2900   2938       LDL-RECEPTOR CLASS A 20.
FT   DOMAIN     2939   2978       EGF-LIKE 11.
FT   DOMAIN     2979   3019       EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
FT   DOMAIN     3287   3328       EGF-LIKE 13.
FT   DOMAIN     3329   3368       LDL-RECEPTOR CLASS A 21.
FT   DOMAIN     3369   3407       LDL-RECEPTOR CLASS A 22.
FT   DOMAIN     3408   3447       LDL-RECEPTOR CLASS A 23.
FT   DOMAIN     3448   3488       LDL-RECEPTOR CLASS A 24.
FT   DOMAIN     3489   3530       LDL-RECEPTOR CLASS A 25.
FT   DOMAIN     3531   3569       LDL-RECEPTOR CLASS A 26.
FT   DOMAIN     3570   3608       LDL-RECEPTOR CLASS A 27.
FT   DOMAIN     3608   3646       LDL-RECEPTOR CLASS A 28.
FT   DOMAIN     3649   3689       LDL-RECEPTOR CLASS A 29.
FT   DOMAIN     3690   3730       LDL-RECEPTOR CLASS A 30.
FT   DOMAIN     3736   3776       LDL-RECEPTOR CLASS A 31.
FT   DOMAIN     3779   3821       EGF-LIKE 14.
FT   DOMAIN     3822   3859       EGF-LIKE 15.
FT   DOMAIN     4146   4182       EGF-LIKE 16.
FT   DOMAIN     4195   4231       EGF-LIKE 17.
FT   DOMAIN     4231   4267       EGF-LIKE 18.
FT   DOMAIN     4267   4303       EGF-LIKE 19.
FT   DOMAIN     4303   4339       EGF-LIKE 20.
FT   DOMAIN     4339   4374       EGF-LIKE 21.
FT   DOMAIN     4372   4409       EGF-LIKE 22.
FT   SITE       3939   3942       RECOGNITION SITE FOR PROTEOLYTICAL
FT                                PROCESSING (POTENTIAL).
```

```
FT   SITE       4501   4506     ENDOCYTOSIS SIGNAL (POTENTIAL).
FT   DISULFID     29     42     BY SIMILARITY.
FT   DISULFID     36     55     BY SIMILARITY.
FT   DISULFID     49     66     BY SIMILARITY.
FT   DISULFID     74     87     BY SIMILARITY.
FT   DISULFID     81    100     BY SIMILARITY.
FT   DISULFID     94    110     BY SIMILARITY.
FT   DISULFID    117    126     BY SIMILARITY.
FT   DISULFID    122    135     BY SIMILARITY.
FT   DISULFID    137    150     BY SIMILARITY.
FT   DISULFID    156    166     BY SIMILARITY.
FT   DISULFID    162    175     BY SIMILARITY.
FT   DISULFID    177    190     BY SIMILARITY.
FT   DISULFID    480    495     BY SIMILARITY.
FT   DISULFID    491    506     BY SIMILARITY.
FT   DISULFID    508    521     BY SIMILARITY.
FT   DISULFID    805    816     BY SIMILARITY.
FT   DISULFID    812    825     BY SIMILARITY.
FT   DISULFID    827    840     BY SIMILARITY.
FT   DISULFID    852    864     BY SIMILARITY.
FT   DISULFID    859    877     BY SIMILARITY.
FT   DISULFID    871    888     BY SIMILARITY.
FT   DISULFID    893    905     BY SIMILARITY.
FT   DISULFID    900    918     BY SIMILARITY.
FT   DISULFID    912    929     BY SIMILARITY.
FT   DISULFID    934    946     BY SIMILARITY.
FT   DISULFID    941    959     BY SIMILARITY.
FT   DISULFID    953    969     BY SIMILARITY.
FT   DISULFID    974    987     BY SIMILARITY.
FT   DISULFID    982   1000     BY SIMILARITY.
FT   DISULFID    994   1009     BY SIMILARITY.
FT   DISULFID   1013   1025     BY SIMILARITY.
FT   DISULFID   1020   1038     BY SIMILARITY.
FT   DISULFID   1032   1049     BY SIMILARITY.
FT   DISULFID   1060   1073     BY SIMILARITY.
FT   DISULFID   1067   1086     BY SIMILARITY.
FT   DISULFID   1080   1095     BY SIMILARITY.
FT   DISULFID   1102   1116     BY SIMILARITY.
FT   DISULFID   1110   1129     BY SIMILARITY.
FT   DISULFID   1123   1138     BY SIMILARITY.
FT   DISULFID   1143   1157     BY SIMILARITY.
FT   DISULFID   1150   1170     BY SIMILARITY.
FT   DISULFID   1164   1180     BY SIMILARITY.
FT   DISULFID   1183   1194     BY SIMILARITY.
FT   DISULFID   1190   1204     BY SIMILARITY.
FT   DISULFID   1206   1219     BY SIMILARITY.
FT   DISULFID   1225   1235     BY SIMILARITY.
FT   DISULFID   1231   1244     BY SIMILARITY.
FT   DISULFID   1246   1259     BY SIMILARITY.
FT   DISULFID   1538   1551     BY SIMILARITY.
FT   DISULFID   1547   1561     BY SIMILARITY.
FT   DISULFID   1563   1576     BY SIMILARITY.
FT   DISULFID   1846   1857     BY SIMILARITY.
FT   DISULFID   1853   1867     BY SIMILARITY.
FT   DISULFID   1869   1882     BY SIMILARITY.
FT   DISULFID   2155   2166     BY SIMILARITY.
FT   DISULFID   2162   2176     BY SIMILARITY.
FT   DISULFID   2178   2190     BY SIMILARITY.
FT   DISULFID   2476   2487     BY SIMILARITY.
FT   DISULFID   2483   2497     BY SIMILARITY.
FT   DISULFID   2499   2511     BY SIMILARITY.
FT   DISULFID   2518   2531     BY SIMILARITY.
FT   DISULFID   2526   2544     BY SIMILARITY.
FT   DISULFID   2538   2555     BY SIMILARITY.
FT   DISULFID   2560   2572     BY SIMILARITY.
FT   DISULFID   2567   2585     BY SIMILARITY.

  Query Match             2.5%;  Score 121;  DB 1;  Length 4543;
  Best Local Similarity   18.7%;  Pred. No. 8;
  Matches  127;  Conservative   73;  Mismatches  190;  Indels  288;  Gaps   38;

Qy         58 CFDASFRGLENCR-----CD----VACKDRGDC---CW------DFED----------- 87
              | |:|     :||      ||      |||   |     |      | ||
Db       1086 CMDSSDE--KNCEGVTHVCDPNVKFGCKDSARCISKAWVCDGDSDCEDNSDEENCESLVC 1143

Qy         88 -----TCVESTRIW-----MCN-KFRCGETRLEASLCSCSDDCLQKKDCCADYKSVCQGE 136
                   ||  :|  |       :|:    ||:   |   ||   | |       |:   :|  ||
Db       1144 KPPSHTCANNTSICLPPEKLCDGSDDCGDGSDEGELC---DQCSLNNGGCSHNCTVAPGE 1200

Qy        137 ----------------TSWLEENC-------DTAQQSQ------CPEGFDL-------- 158
                              |  ::  |         :|  :      | ||:  |
Db       1201 GIVCSCPLGMELGADNKTCQIQSYCAKHLKCSQKCEQDKYNVKCSCYEGWMLEPDGESCR 1260

Qy        159 -----PPVILFS-------MDGFRAEYLYTWDTLMPNINKLKTCGIHSKYMRAMYPTKTF 206
                   | |:||        :|  |  :|        |:|  :          |      ::|  |
Db       1261 SLDPFKPFIIFSNRHEIRRIDLHRGDY----SVLVPGLRNTIALDFHLN-QSSLYWT--- 1312

Qy        207 PNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSKE---QNNPAWWHGQPM-WLTAMYQ 262
                                     :::: :|    |  :|  :|:|  |      |     |    :  |:
Db       1313 --------------DVVEDKIYRGKLLENGALTSFEVVIQYGLATPEGLAVDWI----- 1352

Qy        263 GLKAATYFWPGS-----EVA-INGSFPSIYMPYNGSVPFEERIST-----LLKWLD---- 307
                 |   :|   |      ||| ::|:   :   :     |  :        |:          :|  |  |
Db       1353 ---AGNIYWVESNLDQIEVAKLDGTMRTTLLA--GDIEHPRAIALDPRYGILFWTDWDAS 1407

Qy        308 LPKAERPRF-----YTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLH 362
              ||: |          |:: |      :|  |  |                 |: :: |::|
Db       1408 LPRIEAASMSGAGRRTIHKE----TGSGGWP--------------NGLTVDYLEKR--- 1445

Qy        363 NCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHN-IPHDFFSFNSE 421
                   ||   |    |  |                           :   |     |   :|  |   :  |  |
Db       1446 -----ILWIDARSDAIY-------------SALYDGTGHIEVLRGHEYLSHPF------ 1480

Qy        422 EIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGGG 481
                                                   | |:                  |
Db       1481 ----------------------------------AVTLY-----------------G 1486

Qy        482 NHGYNNEFRSMEAI----FLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHG 537
                   |   ::|:        :    |    :   ::|    :||:  ::||:       |     |    ||   |
Db       1487 GEVYWTDWRTNTLAKANKWTGHNVTVVQRTNTQPFD-LQVYHPSRQPLAPNPCEANGGKG 1545

Qy        538 SLNHLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHL-----QNSTQLEQVNQM 592
                 :||  : :                                :|  |  ||||       |:|   |     :
Db       1546 PCSHLCLINY---------------------NRTLSCACPHLMKLDKDNTTCYEFKKFL 1583

Qy        593 LNLTQEEITATVKVNLPF 610
               |   | ||   |  :: |:
Db       1584 LYARQMEIRG-VDIDNPY 1600

Search completed: July  6, 2004, 13:18:47
Job time : 21 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 6, 2004, 12:45:12 ; Search time 64 Seconds
(without alignments)
3862.954 Million cell updates/sec

Title: US-10-005-480A-743
Perfect score: 4804
Sequence: 1 MESTLTLATEQPVKKNTLKK.....QPVSEILQLKTYLPTFETTI 875

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4804	100.0	875	7	ADE56099	Ade56099 Human Pro
2	4804	100.0	875	7	ADE56103	Ade56103 Human Pro
3	4804	100.0	875	7	ADE65891	Ade65891 Human 161
4	4804	100.0	875	7	ADE65904	Ade65904 Human 161
5	4804	100.0	875	7	ADE65969	Ade65969 Human 161
6	4804	100.0	875	7	ADE65899	Ade65899 Human 161
7	4804	100.0	875	7	ADE65983	Ade65983 Human 161
8	4804	100.0	875	7	ADE65982	Ade65982 Human 161
9	4804	100.0	875	7	ADE65909	Ade65909 Human 161
10	4804	100.0	875	7	ADE65984	Ade65984 Human 161
11	4804	100.0	875	7	ADE65991	Ade65991 Human 161
12	4801	99.9	875	7	ADE65910	Ade65910 Human 161
13	4801	99.9	875	7	ADE65905	Ade65905 Human 161
14	4801	99.9	875	7	ADE65893	Ade65893 Human 161
15	4798	99.9	875	7	ADE65897	Ade65897 Human 161
16	4798	99.9	875	7	ADE65907	Ade65907 Human 161
17	4797	99.9	875	7	ADE65906	Ade65906 Human 161
18	4797	99.9	875	7	ADE65895	Ade65895 Human 161
19	4796	99.8	875	7	ADE65901	Ade65901 Human 161
20	4795	99.8	875	6	ABG74664	Abg74664 Human ect
21	4743	98.7	885	4	ABG10289	Abg10289 Novel hum
22	4642	96.6	841	7	ADE65989	Ade65989 Human 161
23	4642	96.6	841	7	ADE65903	Ade65903 Human 161
24	4634	96.5	841	7	ADE65990	Ade65990 Human 161
25	4634	96.5	841	7	ADE65988	Ade65988 Human 161
26	4634	96.5	841	7	ADE65908	Ade65908 Human 161
27	3953.5	82.3	876	7	ADE56097	Ade56097 Rat Prote
28	3953.5	82.3	876	7	ADE56101	Ade56101 Rat Prote
29	2498.5	52.0	873	3	AAY83620	Aay83620 Human PC-
30	2498.5	52.0	925	2	AAR79148	Aar79148 Human ins
31	2498.5	52.0	925	2	AAY39355	Aay39355 Insulin r
32	2498.5	52.0	925	3	AAB00195	Aab00195 Breast ca
33	2498.5	52.0	925	7	ADD46549	Add46549 Human Pro
34	2494.5	51.9	873	3	AAY83621	Aay83621 Variant h
35	2265	47.1	945	4	AAM25587	Aam25587 Human pro
36	2122.5	44.2	858	4	AAY71998	Aay71998 Rat autot
37	2120	44.1	862	5	ABG32516	Abg32516 Rat lysop
38	2119.5	44.1	858	4	AAY71986	Aay71986 Rat autot
39	2119.5	44.1	858	4	AAY71995	Aay71995 Rat autot
40	2118.5	44.1	858	4	AAY71999	Aay71999 Rat autot
41	2118.5	44.1	858	4	AAY71997	Aay71997 Rat autot
42	2116.5	44.1	858	4	AAY71996	Aay71996 Rat autot
43	2094.5	43.6	863	4	AAY71988	Aay71988 Human ter
44	2094.5	43.6	863	7	ADC10214	Adc10214 Human NOV
45	2091.5	43.5	870	7	ADC10224	Adc10224 Human NOV

ALIGNMENTS

```
RESULT 1
ADE56099
ID   ADE56099 standard; protein; 875 AA.
XX
AC   ADE56099;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human Protein AAC51813, SEQ ID NO 1940.
XX
KW   Human; pain; neuronal tissue; gene therapy;
KW   spinal segmental nerve injury; chronic constriction injury; CCI;
KW   spared nerve injury; SNI; Chung.
XX
OS   Homo sapiens.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (FARB ) BAYER AG.
XX
PI   Woolf C,  D'urso D,  Befort K,  Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; AAC51813.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
```

CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 875 AA;

Query Match 100.0%; Score 4804; DB 7; Length 875;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
```

RESULT 2
ADE56103
ID ADE56103 standard; protein; 875 AA.
XX
AC ADE56103;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein AAC51813, SEQ ID NO 1944.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; AAC51813.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 875 AA;

Query Match 100.0%; Score 4804; DB 7; Length 875;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
```

RESULT 3
ADE65891
ID ADE65891 standard; protein; 875 AA.
XX
AC ADE65891;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human 161P2F10B protein variant 1.
XX
KW 161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.
XX
OS Homo sapiens.
XX
PN WO2003040340-A2.
XX
PD 15-MAY-2003.
XX
PF 07-NOV-2002; 2002WO-US036002.
XX
PR 07-NOV-2001; 2001US-00005480.
PR 31-JAN-2002; 2002US-00062109.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Jakobovits A, Raitano AB, Faris M, Hubert RS, Ge W, Morrison KJM;
PI Morrison RK, Challita-Eid PM;
XX
DR WPI; 2003-441560/41.
DR N-PSDB; ADE66890.
XX
PT A composition for diagnosing, preventing and treating cancer (e.g.
PT prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides
PT and polypeptides.
XX
PS Claim 1; SEQ ID NO 3; 135pp; English.
XX
CC This invention relates to a novel composition which comprises a substance
CC that modulates the status of a novel protein (161P2F10B) and its variants
CC having a sequence of 875 amino acids provided in the specification. The
CC protein of the invention is over-expressed in certain cancers. The
CC compounds of the invention may have cytostatic activity and the sequence
CC of the 161P2F10B protein, and the gene which encodes it, may be useful
CC for gene therapy or the development of a vaccine. The composition and
CC methods of the invention are useful in diagnosing, preventing and
CC treating cancer. The present sequence is the amino acid sequence of the
CC human 161P2F10B variant 1 protein of the invention.
XX
SQ Sequence 875 AA;

Query Match 100.0%; Score 4804; DB 7; Length 875;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


RESULT 4
ADE65904
ID   ADE65904 standard; protein; 875 AA.
XX
AC   ADE65904;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human 161P2F10B protein variant 1 #2.
XX
KW   161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.
XX
OS   Homo sapiens.
XX
PN   WO2003040340-A2.
XX
PD   15-MAY-2003.
XX
PF   07-NOV-2002; 2002WO-US036002.
XX
PR   07-NOV-2001; 2001US-00005480.
PR   31-JAN-2002; 2002US-00062109.
XX
PA   (AGEN-) AGENSYS INC.
XX
PI   Jakobovits A,  Raitano AB,  Paris M,  Hubert RS,  Ge W,  Morrison KJM;
PI   Morrison RK,  Challita-Eid PM;
XX
DR   WPI; 2003-441560/41.
DR   N-PSDB; ADE65891.
XX
PT   A composition for diagnosing, preventing and treating cancer (e.g.
PT   prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides
PT   and polypeptides.
XX
PS   Example 1; SEQ ID NO 16; 135pp; English.
XX
CC   This invention relates to a novel composition which comprises a substance
CC   that modulates the status of a novel protein (161P2F10B) and its variants
CC   having a sequence of 875 amino acids provided in the specification. The
CC   protein of the invention is over-expressed in certain cancers. The
CC   compounds of the invention may have cytostatic activity and the sequence
CC   of the 161P2F10B protein, and the gene which encodes it, may be useful
CC   for gene therapy or the development of a vaccine. The composition and
CC   methods of the invention are useful in diagnosing, preventing and
CC   treating cancer. The present sequence is the amino acid sequence of the
CC   human 161P2F10B variant 1 protein of the invention.
XX
SQ   Sequence 875 AA;

  Query Match           100.0%;  Score 4804;  DB 7;  Length 875;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
                ||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


RESULT 5
ADE65969
ID   ADE65969 standard; protein; 875 AA.
XX
AC   ADE65969;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human 161P2F10B protein variant 1 #4.
XX
KW   161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.
XX
OS   Homo sapiens.
XX
PN   WO2003040340-A2.
XX
PD   15-MAY-2003.
XX
PF   07-NOV-2002; 2002WO-US036002.
XX
PR   07-NOV-2001; 2001US-00005480.
PR   31-JAN-2002; 2002US-00062109.
XX
PA   (AGEN-) AGENSYS INC.
XX
PI   Jakobovits A,  Raitano AB,  Faris M,  Hubert RS,  Ge W,  Morrison KJM;
PI   Morrison RK,  Challita-Eid PM;
XX
DR   WPI; 2003-441560/41.
XX
PT   A composition for diagnosing, preventing and treating cancer (e.g.
PT   prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides
PT   and polypeptides.
XX
PS   Example 65; SEQ ID NO 81; 135pp; English.
XX
CC   This invention relates to a novel composition which comprises a substance
CC   that modulates the status of a novel protein (161P2F10B) and its variants
CC   having a sequence of 875 amino acids provided in the specification. The
CC   protein of the invention is over-expressed in certain cancers. The
CC   compounds of the invention may have cytostatic activity and the sequence
CC   of the 161P2F10B protein, and the gene which encodes it, may be useful
CC   for gene therapy or the development of a vaccine. The composition and
CC   methods of the invention are useful in diagnosing, preventing and
CC   treating cancer. The present sequence is the amino acid sequence of the
CC   human 161P2F10B variant 1 protein of the invention.
XX
SQ   Sequence 875 AA;

  Query Match             100.0%;  Score 4804;  DB 7;  Length 875;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


RESULT 6
ADE65899
ID   ADE65899 standard; protein; 875 AA.
XX
AC   ADE65899;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human 161P2F10B protein variant 5.
XX
KW   161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.
XX
OS   Homo sapiens.
XX
PN   WO2003040340-A2.
XX
PD   15-MAY-2003.
XX
PF   07-NOV-2002; 2002WO-US036002.
XX
```

```
PR   07-NOV-2001; 2001US-00005480.
PR   31-JAN-2002; 2002US-00062109.
XX
PA   (AGEN-) AGENSYS INC.
XX
PI   Jakobovits A,  Raitano AB,  Faris M,  Hubert RS,  Ge W,  Morrison KJM;
PI   Morrison RK,  Challita-Eid PM;
XX
DR   WPI; 2003-441560/41.
DR   N-PSDB; ADE65898.
XX
PT   A composition for diagnosing, preventing and treating cancer (e.g.
PT   prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides
PT   and polypeptides.
XX
PS   Claim 1; SEQ ID NO 11; 135pp; English.
XX
CC   This invention relates to a novel composition which comprises a substance
CC   that modulates the status of a novel protein (161P2F10B) and its variants
CC   having a sequence of 875 amino acids provided in the specification. The
CC   protein of the invention is over-expressed in certain cancers. The
CC   compounds of the invention may have cytostatic activity and the sequence
CC   of the 161P2F10B protein, and the gene which encodes it, may be useful
CC   for gene therapy or the development of a vaccine. The composition and
CC   methods of the invention are useful in diagnosing, preventing and
CC   treating cancer. The present sequence is the amino acid sequence of the
CC   human 161P2F10B variant 5 protein of the invention.
XX
SQ   Sequence 875 AA;

  Query Match           100.0%;  Score 4804;  DB 7;  Length 875;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


RESULT 7
ADE65983
ID   ADE65983 standard; protein; 875 AA.
XX
AC   ADE65983;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human 161P2F10B protein variant 1 #5.
XX
KW   161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.
XX
OS   Homo sapiens.
XX
PN   WO2003040340-A2.
XX
PD   15-MAY-2003.
XX
PF   07-NOV-2002; 2002WO-US036002.
XX
PR   07-NOV-2001; 2001US-00005480.
PR   31-JAN-2002; 2002US-00062109.
XX
PA   (AGEN-) AGENSYS INC.
XX
PI   Jakobovits A,  Raitano AB,  Faris M,  Hubert RS,  Ge W,  Morrison KJM;
PI   Morrison RK,  Challita-Eid PM;
XX
DR   WPI; 2003-441560/41.
XX
PT   A composition for diagnosing, preventing and treating cancer (e.g.
PT   prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides
PT   and polypeptides.
XX
PS   Example 5; SEQ ID NO 95; 135pp; English.
XX
CC   This invention relates to a novel composition which comprises a substance
CC   that modulates the status of a novel protein (161P2F10B) and its variants
CC   having a sequence of 875 amino acids provided in the specification. The
CC   protein of the invention is over-expressed in certain cancers. The
CC   compounds of the invention may have cytostatic activity and the sequence
CC   of the 161P2F10B protein, and the gene which encodes it, may be useful
CC   for gene therapy or the development of a vaccine. The composition and
CC   methods of the invention are useful in diagnosing, preventing and
CC   treating cancer. The present sequence is the amino acid sequence of the
CC   human 161P2F10B variant 1 protein of the invention.
XX
SQ   Sequence 875 AA;

  Query Match           100.0%;  Score 4804;  DB 7;  Length 875;
  Best Local Similarity  100.0%;  Pred. No. 0;
```

```
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


RESULT 8
ADE65982
ID   ADE65982 standard; protein; 875 AA.
XX
AC   ADE65982;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human 161P2F10B protein variant 6 #2.
XX
KW   161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.
XX
OS   Homo sapiens.
XX
PN   WO2003040340-A2.
XX
PD   15-MAY-2003.
XX
PF   07-NOV-2002; 2002WO-US036002.
XX
PR   07-NOV-2001; 2001US-00005480.
PR   31-JAN-2002; 2002US-00062109.
XX
PA   (AGEN-) AGENSYS INC.
XX
PI   Jakobovits A,  Raitano AB,  Faris M,  Hubert RS,  Ge W,  Morrison KJM;
PI   Morrison RK,  Challita-Eid PM;
XX
DR   WPI; 2003-441560/41.
XX
PT   A composition for diagnosing, preventing and treating cancer (e.g.
PT   prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides
PT   and polypeptides.
XX
PS   Example 5; SEQ ID NO 94; 135pp; English.
XX
CC   This invention relates to a novel composition which comprises a substance
CC   that modulates the status of a novel protein (161P2F10B) and its variants
CC   having a sequence of 875 amino acids provided in the specification. The
CC   protein of the invention is over-expressed in certain cancers. The
CC   compounds of the invention may have cytostatic activity and the sequence
CC   of the 161P2F10B protein, and the gene which encodes it, may be useful
CC   for gene therapy or the development of a vaccine. The composition and
CC   methods of the invention are useful in diagnosing, preventing and
CC   treating cancer. The present sequence is the amino acid sequence of the
CC   human 161P2F10B variant 6 protein of the invention.
XX
SQ   Sequence 875 AA;

  Query Match             100.0%;  Score 4804;  DB 7;  Length 875;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
```

```
RESULT 9
ADE65909
ID   ADE65909 standard; protein; 875 AA.
XX
AC   ADE65909;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human 161P2F10B protein variant 1 #3.
XX
KW   161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.
XX
OS   Homo sapiens.
XX
PN   WO2003040340-A2.
XX
PD   15-MAY-2003.
XX
PF   07-NOV-2002; 2002WO-US036002.
XX
PR   07-NOV-2001; 2001US-00005480.
PR   31-JAN-2002; 2002US-00062109.
XX
PA   (AGEN-) AGENSYS INC.
XX
PI   Jakobovits A,  Raitano AB,  Faris M,  Hubert RS,  Ge W,  Morrison KJM;
PI   Morrison RK,  Challita-Eid PM;
XX
DR   WPI; 2003-441560/41.
XX
PT   A composition for diagnosing, preventing and treating cancer (e.g.
PT   prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides
PT   and polypeptides.
XX
PS   Example 2; SEQ ID NO 21; 135pp; English.
XX
CC   This invention relates to a novel composition which comprises a substance
CC   that modulates the status of a novel protein (161P2F10B) and its variants
CC   having a sequence of 875 amino acids provided in the specification. The
CC   protein of the invention is over-expressed in certain cancers. The
CC   compounds of the invention may have cytostatic activity and the sequence
CC   of the 161P2F10B protein, and the gene which encodes it, may be useful
CC   for gene therapy or the development of a vaccine. The composition and
CC   methods of the invention are useful in diagnosing, preventing and
CC   treating cancer. The present sequence is the amino acid sequence of the
CC   human 161P2F10B variant 1 protein of the invention.
XX
SQ   Sequence 875 AA;
```

```
  Query Match           100.0%;  Score 4804;  DB 7;  Length 875;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
```

RESULT 10

```
ADE65984
ID   ADE65984 standard; protein; 875 AA.
XX
AC   ADE65984;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human 161P2F10B protein variant 6 #3.
XX
KW   161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.
XX
OS   Homo sapiens.
XX
PN   WO2003040340-A2.
XX
PD   15-MAY-2003.
XX
PF   07-NOV-2002; 2002WO-US036002.
XX
PR   07-NOV-2001; 2001US-00005480.
PR   31-JAN-2002; 2002US-00062109.
XX
PA   (AGEN-) AGENSYS INC.
XX
PI   Jakobovits A,  Raitano AB,  Faris M,  Hubert RS,  Ge W,  Morrison KJM;
PI   Morrison RK,  Challita-Eid PM;
XX
DR   WPI; 2003-441560/41.
XX
PT   A composition for diagnosing, preventing and treating cancer (e.g.
PT   prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides
PT   and polypeptides.
XX
PS   Example 5; SEQ ID NO 96; 135pp; English.
XX
CC   This invention relates to a novel composition which comprises a substance
CC   that modulates the status of a novel protein (161P2F10B) and its variants
CC   having a sequence of 875 amino acids provided in the specification. The
CC   protein of the invention is over-expressed in certain cancers. The
CC   compounds of the invention may have cytostatic activity and the sequence
CC   of the 161P2F10B protein, and the gene which encodes it, may be useful
CC   for gene therapy or the development of a vaccine. The composition and
CC   methods of the invention are useful in diagnosing, preventing and
CC   treating cancer. The present sequence is the amino acid sequence of the
CC   human 161P2F10B variant 6 protein of the invention.
XX
SQ   Sequence 875 AA;

  Query Match            100.0%;  Score 4804;  DB 7;  Length 875;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


RESULT 11
ADE65991
ID   ADE65991 standard; protein; 875 AA.
XX
AC   ADE65991;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human 161P2F10B protein variant 1 #6.
XX
KW   161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.
XX
OS   Homo sapiens.
XX
PN   WO2003040340-A2.
XX
PD   15-MAY-2003.
XX
PF   07-NOV-2002; 2002WO-US036002.
XX
PR   07-NOV-2001; 2001US-00005480.
PR   31-JAN-2002; 2002US-00062109.
XX
PA   (AGEN-) AGENSYS INC.
XX
PI   Jakobovits A,  Raitano AB,  Faris M,  Hubert RS,  Ge W,  Morrison KJM;
PI   Morrison RK,  Challita-Eid PM;
XX
DR   WPI; 2003-441560/41.
XX
PT   A composition for diagnosing, preventing and treating cancer (e.g.
PT   prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides
PT   and polypeptides.
```

```
XX
PS   Disclosure; Fig 13A; 135pp; English.
XX
CC   This invention relates to a novel composition which comprises a substance
CC   that modulates the status of a novel protein (161P2F10B) and its variants
CC   having a sequence of 875 amino acids provided in the specification. The
CC   protein of the invention is over-expressed in certain cancers. The
CC   compounds of the invention may have cytostatic activity and the sequence
CC   of the 161P2F10B protein, and the gene which encodes it, may be useful
CC   for gene therapy or the development of a vaccine. The composition and
CC   methods of the invention are useful in diagnosing, preventing and
CC   treating cancer. The present sequence is the amino acid sequence of the
CC   human 161P2F10B variant 1 protein of the invention.
XX
SQ   Sequence 875 AA;

  Query Match             100.0%;  Score 4804;  DB 7;  Length 875;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              |||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


RESULT 12
ADE65910
ID   ADE65910 standard; protein; 875 AA.
XX
AC   ADE65910;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human 161P2F10B protein K122R variant.
XX
KW   161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.
XX
OS   Homo sapiens.
XX
PN   WO2003040340-A2.
XX
PD   15-MAY-2003.
XX
PF   07-NOV-2002; 2002WO-US036002.
XX
PR   07-NOV-2001; 2001US-00005480.
PR   31-JAN-2002; 2002US-00062109.
XX
PA   (AGEN-) AGENSYS INC.
XX
PI   Jakobovits A,  Raitano AB,  Faris M,  Hubert RS,  Ge W,  Morrison KJM;
PI   Morrison RK,  Challita-Eid PM;
XX
DR   WPI; 2003-441560/41.
XX
PT   A composition for diagnosing, preventing and treating cancer (e.g.
PT   prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides
PT   and polypeptides.
XX
PS   Example 2; SEQ ID NO 22; 135pp; English.
XX
CC   This invention relates to a novel composition which comprises a substance
CC   that modulates the status of a novel protein (161P2F10B) and its variants
CC   having a sequence of 875 amino acids provided in the specification. The
CC   protein of the invention is over-expressed in certain cancers. The
CC   compounds of the invention may have cytostatic activity and the sequence
CC   of the 161P2F10B protein, and the gene which encodes it, may be useful
CC   for gene therapy or the development of a vaccine. The composition and
CC   methods of the invention are useful in diagnosing, preventing and
CC   treating cancer. The present sequence is the amino acid sequence of the
CC   human 161P2F10B K122R variant protein which was used for homology
CC   purposes and is related to the invention.
XX
SQ   Sequence 875 AA;

  Query Match             99.9%;  Score 4801;  DB 7;  Length 875;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches  874;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QRKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
```

```
Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHPDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
```

```
RESULT 13
ADE65905
ID   ADE65905 standard; protein; 875 AA.
XX
AC   ADE65905;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Human 161P2F10B protein variant 2 #2.
XX
KW   161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.
XX
OS   Homo sapiens.
XX
PN   WO2003040340-A2.
XX
PD   15-MAY-2003.
XX
PF   07-NOV-2002; 2002WO-US036002.
XX
PR   07-NOV-2001; 2001US-00005480.
PR   31-JAN-2002; 2002US-00062109.
XX
PA   (AGEN-) AGENSYS INC.
XX
PI   Jakobovits A,  Raitano AB,  Faris M,  Hubert RS,  Ge W,  Morrison KJM;
PI   Morrison RK,  Challita-Eid PM;
XX
DR   WPI; 2003-441560/41.
DR   N-PSDB; ADE65893.
XX
PT   A composition for diagnosing, preventing and treating cancer (e.g.
PT   prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides
PT   and polypeptides.
XX
PS   Example 1; SEQ ID NO 17; 135pp; English.
XX
CC   This invention relates to a novel composition which comprises a substance
CC   that modulates the status of a novel protein (161P2F10B) and its variants
CC   having a sequence of 875 amino acids provided in the specification. The
CC   protein of the invention is over-expressed in certain cancers. The
CC   compounds of the invention may have cytostatic activity and the sequence
CC   of the 161P2F10B protein, and the gene which encodes it, may be useful
CC   for gene therapy or the development of a vaccine. The composition and
CC   methods of the invention are useful in diagnosing, preventing and
CC   treating cancer. The present sequence is the amino acid sequence of the
CC   human 161P2F10B variant 2 protein of the invention.
XX
SQ   Sequence 875 AA;

  Query Match           99.9%;  Score 4801;  DB 7;  Length 875;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches  874;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QRKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875

Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
```

RESULT 14
ADE65893
ID ADE65893 standard; protein; 875 AA.
XX
AC ADE65893;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human 161P2F10B protein variant 2.
XX
KW 161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.
XX
OS Homo sapiens.
XX
PN WO2003040340-A2.
XX
PD 15-MAY-2003.
XX
PF 07-NOV-2002; 2002WO-US036002.
XX
PR 07-NOV-2001; 2001US-00005480.
PR 31-JAN-2002; 2002US-00062109.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Jakobovits A, Raitano AB, Faris M, Hubert RS, Ge W, Morrison KJM;
PI Morrison RK, Challita-Eid PM;
XX
DR WPI; 2003-441560/41.
DR N-PSDB; ADE65892.
XX
PT A composition for diagnosing, preventing and treating cancer (e.g.
PT prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides
PT and polypeptides.
XX
PS Claim 1; SEQ ID NO 5; 135pp; English.
XX
CC This invention relates to a novel composition which comprises a substance
CC that modulates the status of a novel protein (161P2F10B) and its variants
CC having a sequence of 875 amino acids provided in the specification. The
CC protein of the invention is over-expressed in certain cancers. The
CC compounds of the invention may have cytostatic activity and the sequence
CC of the 161P2F10B protein, and the gene which encodes it, may be useful
CC for gene therapy or the development of a vaccine. The composition and
CC methods of the invention are useful in diagnosing, preventing and
CC treating cancer. The present sequence is the amino acid sequence of the
CC human 161P2F10B variant 2 protein of the invention.
XX
SQ Sequence 875 AA;

Query Match 99.9%; Score 4801; DB 7; Length 875;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 874; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Db        121 QRKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875

Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
```

RESULT 15
ADE65897
ID ADE65897 standard; protein; 875 AA.
XX
AC ADE65897;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human 161P2F10B protein variant 4.
XX
KW 161P2F10B; cancer; cytostatic; gene therapy; vaccine; human.
XX

OS Homo sapiens.
XX
PN WO2003040340-A2.
XX
PD 15-MAY-2003.
XX
PF 07-NOV-2002; 2002WO-US036002.
XX
PR 07-NOV-2001; 2001US-00005480.
PR 31-JAN-2002; 2002US-00062109.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Jakobovits A, Raitano AB, Faris M, Hubert RS, Ge W, Morrison KJM;
PI Morrison RK, Challita-Eid PM;
XX
DR WPI; 2003-441560/41.
DR N-PSDB; ADE65896.
XX
PT A composition for diagnosing, preventing and treating cancer (e.g.
PT prostatic, renal or uterine cancer) comprises 161P2F10B polynucleotides
PT and polypeptides.
XX
PS Claim 1; SEQ ID NO 9; 135pp; English.
XX
CC This invention relates to a novel composition which comprises a substance
CC that modulates the status of a novel protein (161P2F10B) and its variants
CC having a sequence of 875 amino acids provided in the specification. The
CC protein of the invention is over-expressed in certain cancers. The
CC compounds of the invention may have cytostatic activity and the sequence
CC of the 161P2F10B protein, and the gene which encodes it, may be useful
CC for gene therapy or the development of a vaccine. The composition and
CC methods of the invention are useful in diagnosing, preventing and
CC treating cancer. The present sequence is the amino acid sequence of the
CC human 161P2F10B variant 4 protein of the invention.
XX
SQ Sequence 875 AA;

Query Match 99.9%; Score 4798; DB 7; Length 875;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 874; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              |||||||||||||||||||||||||||||||||| |
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETPI 875
```

Search completed: July 6, 2004, 13:18:17
Job time : 68 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 6, 2004, 13:17:02 ; Search time 23 Seconds
(without alignments)
1964.032 Million cell updates/sec

Title: US-10-005-480A-743
Perfect score: 4804
Sequence: 1 MESTLTLATEQPVKKNTLKK.....QPVSEILQLKTYLPTFETTI 875

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2498.5	52.0	873	3	US-09-187-331-6	Sequence 6, Appli
2	2498.5	52.0	873	4	US-09-470-946-6	Sequence 6, Appli
3	2498.5	52.0	873	4	US-09-438-906-2	Sequence 2, Appli
4	2498.5	52.0	925	2	US-08-392-946-1	Sequence 1, Appli
5	2498.5	52.0	925	2	US-08-504-169-1	Sequence 1, Appli
6	2498.5	52.0	925	5	PCT-US94-14893-1	Sequence 1, Appli
7	2494.5	51.9	873	4	US-09-438-906-4	Sequence 4, Appli
8	2040.5	42.5	915	1	US-08-346-455B-69	Sequence 69, Appl
9	2040.5	42.5	915	3	US-08-977-221-69	Sequence 69, Appl
10	2040.5	42.5	915	4	US-09-483-831B-69	Sequence 69, Appl
11	2040.5	42.5	915	5	PCT-US95-06613-69	Sequence 69, Appl
12	1964.5	40.9	829	1	US-08-346-455B-34	Sequence 34, Appl
13	1964.5	40.9	829	3	US-08-977-221-34	Sequence 34, Appl
14	1964.5	40.9	829	4	US-09-483-831B-34	Sequence 34, Appl
15	1964.5	40.9	829	5	PCT-US95-06613-34	Sequence 34, Appl
16	1923.5	40.0	861	1	US-08-346-455B-67	Sequence 67, Appl
17	1923.5	40.0	861	3	US-08-977-221-67	Sequence 67, Appl
18	1923.5	40.0	861	4	US-09-483-831B-67	Sequence 67, Appl
19	1923.5	40.0	861	5	PCT-US95-06613-67	Sequence 67, Appl
20	1838.5	38.3	979	1	US-08-346-455B-38	Sequence 38, Appl
21	1838.5	38.3	979	3	US-08-977-221-38	Sequence 38, Appl
22	1838.5	38.3	979	4	US-09-483-831B-70	Sequence 70, Appl
23	1838.5	38.3	979	5	PCT-US95-06613-38	Sequence 38, Appl
24	1761.5	36.7	788	1	US-08-346-455B-36	Sequence 36, Appl
25	1761.5	36.7	788	3	US-08-977-221-36	Sequence 36, Appl
26	1761.5	36.7	788	4	US-09-483-831B-36	Sequence 36, Appl
27	1761.5	36.7	788	5	PCT-US95-06613-36	Sequence 36, Appl
28	641	13.3	453	4	US-09-800-729-83	Sequence 83, Appl
29	503	10.5	355	2	US-09-014-969-19	Sequence 19, Appl
30	432.5	9.0	438	3	US-09-187-331-2	Sequence 2, Appli
31	432.5	9.0	438	4	US-09-470-946-2	Sequence 2, Appli
32	227.5	4.7	136	4	US-09-621-976-3913	Sequence 3913, Ap
33	210.5	4.4	151	4	US-09-621-976-3891	Sequence 3891, Ap
34	144.5	3.0	1354	4	US-07-757-022B-48	Sequence 48, Appl
35	143	3.0	1022	4	US-07-757-022B-84	Sequence 84, Appl
36	140.5	2.9	237	4	US-07-757-022B-72	Sequence 72, Appl
37	139.5	2.9	1311	4	US-07-757-022B-42	Sequence 42, Appl
38	139	2.9	463	4	US-07-757-022B-54	Sequence 54, Appl
39	136.5	2.8	157	4	US-07-757-022B-102	Sequence 102, App
40	136.5	2.8	157	4	US-07-757-022B-114	Sequence 114, App
41	134	2.8	106	4	US-07-757-022B-56	Sequence 56, Appl
42	134	2.8	130	4	US-07-757-022B-78	Sequence 78, Appl
43	134	2.8	130	4	US-07-757-022B-86	Sequence 86, Appl
44	134	2.8	138	4	US-07-757-022B-34	Sequence 34, Appl
45	134	2.8	141	4	US-07-757-022B-80	Sequence 80, Appl

ALIGNMENTS

RESULT 1
US-09-187-331-6
; Sequence 6, Application US/09187331
; Patent No. 6043056
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Gorgone, Gina A.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS
; FILE REFERENCE: PF-0631 US
; CURRENT APPLICATION NUMBER: US/09/187,331
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 873
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: g189650
US-09-187-331-6

Query Match 52.0%; Score 2498.5; DB 3; Length 873;
Best Local Similarity 53.1%; Pred. No. 8e-235;
Matches 465; Conservative 137; Mismatches 250; Indels 23; Gaps 11;

```
Qy         10 EQPVKK----NTLK---KYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFDAS 62
              |:|::|     | |     ||:  :||   ::    ||    ||      ::  ||: :||: :
Db          5 EEPLEKAARARTAKDPNTYKVLSLVLSVCVLTTILGCIFGLKPSCAKEVKSCKGRCFERT 64

Qy         63 FRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQK 122
              |    ||||| || : |:|| |:::||:|    || |||||||| ||  |||:|||||  |
Db         65 F---GNCRCDAACVELGNCCLDYQETCIEPEHIWTCNKFRCGEKRLTRSLCACSDDCKDK 121

Qy        123 KDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMP 182
              ||| :| |||||| ||:|| |::  : ||| ||: || :|||:|||||||||:||  |:|
Db        122 GDCCINYSSVCQGEKSWVEEPCESINEPQCPAGFETPPTLLFSLDGFRAEYLHTWGGLLP 181

Qy        183 NINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSK 242
              |:||| || ::| || :||||||||||:|||||||||||||| |||  :| :||| ||
Db        182 VISKLKKCGTYTKNMRPVYPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMNASFSLKSK 241

Qy        243 EQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTL 302
              |: || |: |:|:|:|| |||||: |:|||||:| ||| || ||  |||||||||| :
Db        242 EKFNPEWYKGEPIWVTAKYQGLKSGTFFWPGSDVEINGIFPDIYKMYNGSVPFEERILAV 301

Qy        303 LKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLH 362
              |:|| ||| ||| |||:| |||||||: |||| |||||| ||   ||||:|||: |||
```

```
Db        302 LQWLQLPKDERPHFYTLYLEEPDSSGHSYGPVSSEVIKALQRVDGMVGMLMDGLKELNLH 361

Qy        363 NCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNSEE 422

Db        362 RCLNLILISDHGMEQGSCKKYIYLNKYLGDVKNIKVIYGPAARLRPSDVPDKYYSFNYEG 421

Qy        423 IVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW-LAVRSKSNTNCGGG 481

Db        422 IARNLSCREPNQHFKPYLKHFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSERKYCGSG 481

Qy        482 NHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNH 541

Db        482 FHGSDNVFSNMQALFVGYGPGFKHGIEADTFENIEVYNLMCDLLNLTPAPNNGTHGSLNH 541

Qy        542 LLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQMLNLTQEEI 600

Db        542 LLKNPVYTPKHPKEVHPLVQCPFTRN-PRDNLGCSCNPSI---LPIEDFQTQPNLTVAEE 597

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Db        598 KIIKHETLPYGRPRVLQKENTICLLSQHQFMSGYSQDILMPLWTSYTVDR--NDSFSTED 655

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQY-DALITSNLVPMYE 719

Db        656 FSNCLYQDFRIPLSPVHKCSFYKNNTKVSYGFLSPPQLNKNSSGIYSEALLTTNIVPMYQ 715

Qy        720 EFRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKH---LANTDVPI 776

Db        716 SFQVIWRYFHDTLLRKYAEERNGVNVVSGPVFDFDYDGRCDSLENLRQKRRVIRNQEILI 775

Qy        777 PTHYFVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHI 836

Db        776 PTHFFIVLTSCKDTSQTPLHCEN-LDTLAFILPHRTDNSESCVHGKHDSSWVEELLMLHR 834

Qy        837 ARVRDVELLTGLDFYQDKVQPVSEILQLKTYLPTF 871

Db        835 ARITDVEHITGLSFYQQRKEPVSDILKLKTHLPTF 869


RESULT 2
US-09-470-946-6
; Sequence 6, Application US/09470946
; Patent No. 6358923
; GENERAL INFORMATION:
;  APPLICANT: Yue, Henry
;  APPLICANT: Corley, Neil C.
;  APPLICANT: Guegler, Karl J.
;  APPLICANT: Gorgone, Gina A.
;  APPLICANT: Baughn, Mariah R.
;  TITLE OF INVENTION: CELL SURFACE GLYCOPROTEINS
;  FILE REFERENCE: PF-0631 US
;  CURRENT APPLICATION NUMBER: US/09/470,946
;  CURRENT FILING DATE: 1999-12-22
;  EARLIER APPLICATION NUMBER: US 09/187,331
;  EARLIER FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: PERL Program
; SEQ ID NO 6
;   LENGTH: 873
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE: -
;   OTHER INFORMATION: g189650
US-09-470-946-6

  Query Match            52.0%;  Score 2498.5;  DB 4;  Length 873;
  Best Local Similarity  53.1%;  Pred. No. 8e-235;
  Matches  465;  Conservative  137;  Mismatches  250;  Indels   23;  Gaps   11;

Qy         10 EQPVKK----NTLK---KYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFDAS 62

Db          5 EEPLEKAARARTAKDPNTYKVLSLVLSVCVLTTILGCIFGLKPSCAKEVKSCKGRCFERT 64

Qy         63 FRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQK 122

Db         65 F---GNCRCDAACVELGNCCLDYQETCIEPEHIWTCNKFRCGEKRLTRSLCACSDDCKDK 121

Qy        123 KDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMP 182

Db        122 GDCCINYSSVCQGEKSWVEEPCESINEPQCPAGFETPPTLLFSLDGFRAEYLHTWGGLLP 181

Qy        183 NINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSK 242

Db        182 VISKLKKCGTYTKNMRPVYPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMNASFSLKSK 241

Qy        243 EQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTL 302

Db        242 EKFNPEWYKGEPIWVTAKYQGLKSGTFFWPGSDVEINGIFPDIYKMYNGSVPFEERILAV 301

Qy        303 LKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLH 362

Db        302 LQWLQLPKDERPHFYTLYLEEPDSSGHSYGPVSSEVIKALQRVDGMVGMLMDGLKELNLH 361

Qy        363 NCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNSEE 422

Db        362 RCLNLILISDHGMEQGSCKKYIYLNKYLGDVKNIKVIYGPAARLRPSDVPDKYYSFNYEG 421

Qy        423 IVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW-LAVRSKSNTNCGGG 481

Db        422 IARNLSCREPNQHFKPYLKHFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSERKYCGSG 481

Qy        482 NHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNH 541

Db        482 FHGSDNVFSNMQALFVGYGPGFKHGIEADTFENIEVYNLMCDLLNLTPAPNNGTHGSLNH 541

Qy        542 LLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQMLNLTQEEI 600

Db        542 LLKNPVYTPKHPKEVHPLVQCPFTRN-PRDNLGCSCNPSI---LPIEDFQTQFNLTVAEE 597

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Db        598 KIIKHETLPYGRPRVLQKENTICLLSQHQFMSGYSQDILMPLWTSYTVDR--NDSFSTED 655

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQY-DALITSNLVPMYE 719

Db        656 FSNCLYQDFRIPLSPVHKCSFYKNNTKVSYGFLSPPQLNKNSSGIYSEALLTTNIVPMYQ 715

Qy        720 EFRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKH---LANTDVPI 776

Db        716 SFQVIWRYFHDTLLRKYAEERNGVNVVSGPVFDFDYDGRCDSLENLRQKRRVIRNQEILI 775

Qy        777 PTHYFVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHI 836

Db        776 PTHFFIVLTSCKDTSQTPLHCEN-LDTLAFILPHRTDNSESCVHGKHDSSWVEELLMLHR 834

Qy        837 ARVRDVELLTGLDFYQDKVQPVSEILQLKTYLPTF 871

Db        835 ARITDVEHITGLSFYQQRKEPVSDILKLKTHLPTF 869


RESULT 3
US-09-438-906-2
; Sequence 2, Application US/09438906
; Patent No. 6465185
; GENERAL INFORMATION:
;  APPLICANT: Goldfine, Ira
;  APPLICANT:  Trischitta, Vincenzo
;  APPLICANT:  Pizzuti, Antonio
;  APPLICANT:  Vigneri, Riccardo
;  APPLICANT:  Frittitta, Lucia
;  TITLE OF INVENTION: Polymorphic Human PC-1 Sequences
;  FILE REFERENCE: 9076-089
;  CURRENT APPLICATION NUMBER: US/09/438,906
;  CURRENT FILING DATE:  1999-11-18
;  PRIOR APPLICATION NUMBER: 60/108,853
;  PRIOR FILING DATE: 1998-11-18
```

```
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 873
;   TYPE: PRT
;   ORGANISM: H. sapiens
US-09-438-906-2

  Query Match             52.0%;  Score 2498.5;  DB 4;  Length 873;
  Best Local Similarity   53.1%;  Pred. No. 8e-235;
  Matches  465;  Conservative  137;  Mismatches  250;  Indels   23;  Gaps   11;

Qy         10 EQPVKK----NTLK---KYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFDAS 62
Db          5 EEPLEKAARARTAKDPNTYKVLSLVLSVCVLTTILGCIFGLKPSCAKEVKSCKGRCFERT 64

Qy         63 FRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQK 122
Db         65 F---GNCRCDAACVELGNCCLDYQETCIEPEHIWTCNKFRCGEKRLTRSLCACSDDCKDK 121

Qy        123 KDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMP 182
Db        122 GDCCINYSSVCQGEKSWVEEPCESINEPQCPAGFETPPTLLFSLDGFRAEYLHTWGGLLP 181

Qy        183 NINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSK 242
Db        182 VISKLKKCGTYTKNMRPVYPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMNASFSLKSK 241

Qy        243 EQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTL 302
Db        242 EKFNPEWYKGEPIWVTAKYQGLKSGTFFWPGSDVEINGIFPDIYKMYNGSVPFEERILAV 301

Qy        303 LKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLH 362
Db        302 LQWLQLPKDERPHFYTLYLEEPDSSGHSYGPVSSEVIKALQRVDGMVGMLMDGLKELNLH 361

Qy        363 NCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNSEE 422
Db        362 RCLNLILISDHGMEQGSCKKYIYLNKYLGDVKNIKVIYGPAARLRPSDVPDKYYSFNYEG 421

Qy        423 IVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW-LAVRSKSNTNCGGG 481
Db        422 IARNLSCREPNQHFKPYLKHFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSERKYCGSG 481

Qy        482 NHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNH 541
Db        482 FHGSDNVFSNMQALFVGYGPGFKHGIEADTFENIEVYNLMCDLLNLTPAPNNGTHGSLNH 541

Qy        542 LLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQMLNLTQEEI 600
Db        542 LLKNPVYTPKHPKEVHPLVQCPFTRN-PRDNLGCSCNPSI---LPIEDFQTQFNLTVAEE 597

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
Db        598 KIIKHETLPYGRPRVLQKENTICLLSQHQFMSGYSQDILMPLWTSYTVDR--NDSFSTED 655

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQY-DALITSNLVPMYE 719
Db        656 FSNCLYQDFRIPLSPVHKCSFYKNNTKVSYGFLSPPQLNKNSSGIYSEALLTTNIVPMYQ 715

Qy        720 EFRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKH---LANTDVPI 776
Db        716 SFQVIWRYFHDTLLRKYAEERNGVNVVSGPVFDFDYDGRCDSLENLRQKRRVIRNQEILI 775

Qy        777 PTHYFVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHI 836
Db        776 PTHFFIVLTSCKDTSQTPLHCEN-LDTLAFILPHRTDNSESCVHGKHDSSWVEELLMLHR 834

Qy        837 ARVRDVELLTGLDFYQDKVQPVSEILQLKTYLPTF 871
Db        835 ARITDVEHITGLSFYQQRKEPVSDILKLKTHLPTF 869
```

```
RESULT 4
US-08-392-946-1
; Sequence 1, Application US/08392946
; Patent No. 5939269
;  GENERAL INFORMATION:
;    APPLICANT:  Genentech, Inc.
;    APPLICANT:  The Regents of the University of California
;    APPLICANT:  Goldfine, Ira D.
;    APPLICANT:  Grupe, Andrew
;    APPLICANT:  Maddux, Betty A.
;    APPLICANT:  Spencer, Steven
;    APPLICANT:  Stewart, Timothy A.
;    TITLE OF INVENTION:  Antagonists to Insulin Receptor Tyrosine
;    TITLE OF INVENTION:  Kinase Inhibitor
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  1 DNA Way
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/392,946
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/182241
;      FILING DATE:  14-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kubinec, Jeffrey S.
;      REGISTRATION NUMBER:  36,575
;      REFERENCE/DOCKET NUMBER:  P0875P1PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650/225-8228
;      TELEFAX:  650/952-9881
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  925 amino acids
;      TYPE:  Amino Acid
;      TOPOLOGY:  Linear
US-08-392-946-1

  Query Match             52.0%;  Score 2498.5;  DB 2;  Length 925;
  Best Local Similarity   53.1%;  Pred. No. 8.8e-235;
  Matches  465;  Conservative  137;  Mismatches  250;  Indels   23;  Gaps   11;

Qy         10 EQPVKK----NTLK---KYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFDAS 62
Db         57 EEPLEKAARARTAKDPNTYKVLSLVLSVCVLTTILGCIFGLKPSCAKEVKSCKGRCFERT 116

Qy         63 FRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQK 122
Db        117 F---GNCRCDAACVELGNCCLDYQETCIEPEHIWTCNKFRCGEKRLTRSLCACSDDCKDK 173

Qy        123 KDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMP 182
Db        174 GDCCINYSSVCQGEKSWVEEPCESINEPQCPAGFETPPTLLFSLDGFRAEYLHTWGGLLP 233

Qy        183 NINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSK 242
Db        234 VISKLKKCGTYTKNMRPVYPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMNASFSLKSK 293

Qy        243 EQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTL 302
Db        294 EKFNPEWYKGEPIWVTAKYQGLKSGTFFWPGSDVEINGIFPDIYKMYNGSVPFEERILAV 353

Qy        303 LKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLH 362
```

```
Db        354 LQWLQLPKDERPHFYTLYLEEPDSSGHSYGPVSSEVIKALQRVDGMVGMLMDGLKELNLH 413

Qy        363 NCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNSEE 422

Db        414 RCLNLILISDHGMEQGSCKKYIYLNKYLGDVKNIKVIYGPAARLRPSDVPDKYYSFNYEG 473

Qy        423 IVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW-LAVRSKSNTNCGGG 481

Db        474 IARNLSCREPNQHFKPYLKHFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSERKYCGSG 533

Qy        482 NHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNH 541

Db        534 FHGSDNVFSNMQALFVGYGPGFKHGIEADTFENIEVYNLMCDLLNLTPAPNNGTHGSLNH 593

Qy        542 LLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQMLNLTQEEI 600

Db        594 LLKNPVYTPKHPKEVHPLVQCPFTRN-PRDNLGCSCNPSI---LPIEDFQTQFNLTVAEE 649

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Db        650 KIIKHETLPYGRPRVLQKENTICLLSQHQFMSGYSQDILMPLWTSYTVDR--NDSFSTED 707

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQY-DALITSNLVPMYE 719

Db        708 FSNCLYQDFRIPLSPVHKCSFYKNNTKVSYGFLSPPQLNKNSSGIYSEALLTTNIVPMYQ 767

Qy        720 EFRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKH---LANTDVPI 776

Db        768 SFQVIWRYFHDTLLRKYAEERNGVNVVSGPVFDFDYDGRCDSLENLRQKRRVIRNQEILI 827

Qy        777 PTHYFVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHI 836

Db        828 PTHFFIVLTSCKDTSQTPLHCEN-LDTLAFILPHRTDNSESCVHGKHDSSWVEELLMLHR 886

Qy        837 ARVRDVELLTGLDFYQDKVQPVSEILQLKTYLPTF 871

Db        887 ARITDVEHITGLSFYQQRKEPVSDILKLKTHLPTF 921


RESULT 5
US-08-504-169-1
; Sequence 1, Application US/08504169
; Patent No. 5968508
;  GENERAL INFORMATION:
;    APPLICANT:  Goldfine, Ira
;    APPLICANT:  Grupe, Andrew
;    APPLICANT:  Henzel, William
;    APPLICANT:  Maddox, Betty
;    APPLICANT:  Spencer, Steven
;    APPLICANT:  Stewart, Timothy A.
;    APPLICANT:  Harrison, Denise M.
;    TITLE OF INVENTION:  Antibodies to Insulin Receptor Tyrosine Kinase Activation Inhi
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/504,169
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/14893
;      FILING DATE:  28-Dec-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/182241
;      FILING DATE:  14-Jan-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kubinec, Jeffrey S.
;      REGISTRATION NUMBER:  36,575
;      REFERENCE/DOCKET NUMBER:  P0875P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-8228
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  925 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-504-169-1

  Query Match           52.0%;  Score 2498.5;  DB 2;  Length 925;
  Best Local Similarity  53.1%;  Pred. No. 8.8e-235;
  Matches  465;  Conservative  137;  Mismatches  250;  Indels   23;  Gaps   11;

Qy         10 EQPVKK----NTLK---KYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFDAS 62

Db         57 EEPLEKAARARTAKDPNTYKVLSLVLSVCVLTTILGCIFGLKPSCAKEVKSCKGRCFERT 116

Qy         63 FRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQK 122

Db        117 F---GNCRCDAACVELGNCCLDYQETCIEPEHIWTCNKFRCGEKRLTRSLCACSDDCKDK 173

Qy        123 KDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMP 182

Db        174 GDCCINYSSVCQGEKSWVEEPCESINEPQCPAGFETPPTLLFSLDGFRAEYLHTWGGLLP 233

Qy        183 NINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSK 242

Db        234 VISKLKKCGTYTKNMRPVYPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMNASFSLKSK 293

Qy        243 EQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTL 302

Db        294 EKFNPEWYKGEPIWVTAKYQGLKSGTFFWPGSDVEINGIFPDIYKMYNGSVPFEERILAV 353

Qy        303 LKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLH 362

Db        354 LQWLQLPKDERPHFYTLYLEEPDSSGHSYGPVSSEVIKALQRVDGMVGMLMDGLKELNLH 413

Qy        363 NCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNSEE 422

Db        414 RCLNLILISDHGMEQGSCKKYIYLNKYLGDVKNIKVIYGPAARLRPSDVPDKYYSFNYEG 473

Qy        423 IVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW-LAVRSKSNTNCGGG 481

Db        474 IARNLSCREPNQHFKPYLKHFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSERKYCGSG 533

Qy        482 NHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNH 541

Db        534 FHGSDNVFSNMQALFVGYGPGFKHGIEADTFENIEVYNLMCDLLNLTPAPNNGTHGSLNH 593

Qy        542 LLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQMLNLTQEEI 600

Db        594 LLKNPVYTPKHPKEVHPLVQCPFTRN-PRDNLGCSCNPSI---LPIEDFQTQFNLTVAEE 649

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Db        650 KIIKHETLPYGRPRVLQKENTICLLSQHQFMSGYSQDILMPLWTSYTVDR--NDSFSTED 707

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQY-DALITSNLVPMYE 719

Db        708 FSNCLYQDFRIPLSPVHKCSFYKNNTKVSYGFLSPPQLNKNSSGIYSEALLTTNIVPMYQ 767

Qy        720 EFRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKH---LANTDVPI 776

Db        768 SFQVIWRYFHDTLLRKYAEERNGVNVVSGPVFDFDYDGRCDSLENLRQKRRVIRNQEILI 827
```

```
Qy        777 PTHYFVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHI 836
              |||:|:|||||||: | || :|   || | ||:|||  | |||  || :: ||||    |
Db        828 PTHFFIVLTSCKDTSQTPLHCEN-LDTLAFILPHRTDNSESCVHGKHDSSWVEELLMLHR 886

Qy        837 ARVRDVELLTGLDFYQDKVQPVSEILQLKTYLPTF 871
              ||: ||| :||| ||| : :|||:||:|||:||||
Db        887 ARITDVEHITGLSFYQQRKEPVSDILKLKTHLPTF 921


RESULT 6
PCT-US94-14893-1
; Sequence 1, Application PC/TUS9414893
;  GENERAL INFORMATION:
;    APPLICANT:  Genentech, Inc.
;    APPLICANT:  The Regents of the University of California
;    APPLICANT:  Goldfine, Ira D.
;    APPLICANT:  Grupe, Andrew
;    APPLICANT:  Maddux, Betty A.
;    APPLICANT:  Spencer, Steven
;    APPLICANT:  Stewart, Timothy A.
;    TITLE OF INVENTION:  Antagonists to Insulin Receptor Tyrosine Kinase Inhibitor
;    NUMBER OF SEQUENCES:  1
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/14893
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/182241
;      FILING DATE:  14-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kubinec, Jeffrey S.
;      REGISTRATION NUMBER:  36,575
;      REFERENCE/DOCKET NUMBER:  875P1PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  925 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
PCT-US94-14893-1

  Query Match             52.0%;  Score 2498.5;  DB 5;  Length 925;
  Best Local Similarity   53.1%;  Pred. No. 8.8e-235;
  Matches  465;  Conservative  137;  Mismatches  250;  Indels   23;  Gaps   11;

Qy         10 EQPVKK----NTLK---KYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFDAS 62
              |:|::|      | |    ||:  :||   ::   ||   ||      ::  ||: :||: :
Db         57 EEPLEKAARARTAKDPNTYKVLSLVLSVCVLTTILGCIFGLKPSCAKEVKSCKGRCFERT 116

Qy         63 FRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQK 122
              |     ||||| || : |:|| |:::||:|    || ||||||| ||  |||:|||||  |
Db        117 F---GNCRCDAACVELGNCCLDYQETCIEPEHIWTCNKFRCGEKRLTRSLCACSDDCKDK 173

Qy        123 KDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMP 182
               ||| :| |||||| ||:|| |::  : ||| ||: || :|||:||||||||:||  |:|
Db        174 GDCCINYSSVCQGEKSWVEEPCESINEPQCPAGFETPPTLLFSLDGFRAEYLHTWGGLLP 233

Qy        183 NINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSK 242
              |:||| || ::| || :||||||||||:|||||||||||||| |||  :| :||| ||
Db        234 VISKLKKCGTYTKNMRPVYPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMNASFSLKSK 293

Qy        243 EQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTL 302
              |: || |: |:|:|:|| ||||||: |:|||||:| ||| || ||  |||||||||| :
Db        294 EKFNPEWYKGEPIWVTAKYQGLKSGTFFWPGSDVEINGIFPDIYKMYNGSVPFEERILAV 353

Qy        303 LKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLH 362
              |:|| ||| ||| |||:| |||||||: ||||: |||||  ||   ||||:|||: |||
Db        354 LQWLQLPKDERPHFYTLYLEEPDSSGHSYGPVSSEVIKALQRVDGMVGMLMDGLKELNLH 413

Qy        363 NCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNSEE 422
              |:|:||::|||:|  | |  |:  |   :   :  ||| |:|  ::|  ::||| |
Db        414 RCLNLILISDHGMEQGSCKKYIYLNKYLGDVKNIKVIYGPAARLRPSDVPDKYYSFNYEG 473

Qy        423 IVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW-LAVRSKSNTNCGGG 481
              | ||||||:|:|||||||   ||||||:||: ||: :  ::| || ||:       || |
Db        474 IARNLSCREPNQHFKPYLKHFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSERKYCGSG 533

Qy        482 NHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNH 541
              || :| | :|:|:|: :|| ||   | : |||||||||||||| : ||||||||||||
Db        534 FHGSDNVFSNMQALFVGYGPGFKHGIEADTFENIEVYNLMCDLLNLTPAPNNGTHGSLNH 593

Qy        542 LLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQMLNLTQEEI 600
              ||| | | | | :||     | |    | ::| | | | :     :|      ||| |
Db        594 LLKNPVYTPKHPKEVHPLVQCPFTRN-PRDNLGCSCNPSI---LPIEDFQTQFNLTVAEE 649

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
                    ||:|||||||    |||   :::||: : :||:|:|||| :  : |
Db        650 KIIKHETLPYGRPRVLQKENTICLLSQHQFMSGYSQDILMPLWTSYTVDR--NDSFSTED 707

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQY-DALITSNLVPMYE 719
              :||  | |:| |  |||||  :  :::||| ||  |: |   | :||:|:|:||||:
Db        708 FSNCLYQDFRIPLSPVHKCSFYKNNTKVSYGFLSPPQLNKNSSGIYSEALLTTNIVPMYQ 767

Qy        720 EFRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHPDAPDEITKH---LANTDVPI 776
              |: :| |||  || |:| ||||||||||:||::|||  |: : : :    : | :: |
Db        768 SFQVIWRYFHDTLLRKYAEERNGVNVVSGPVFDFDYDGRCDSLENLRQKRRVIRNQEILI 827

Qy        777 PTHYFVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHI 836
              |||:|:|||||||: | || :|   || | ||:|||  | |||  || :: ||||    |
Db        828 PTHFFIVLTSCKDTSQTPLHCEN-LDTLAFILPHRTDNSESCVHGKHDSSWVEELLMLHR 886

Qy        837 ARVRDVELLTGLDFYQDKVQPVSEILQLKTYLPTF 871
              ||: ||| :||| ||| : :|||:||:|||:||||
Db        887 ARITDVEHITGLSFYQQRKEPVSDILKLKTHLPTF 921


RESULT 7
US-09-438-906-4
; Sequence 4, Application US/09438906
; Patent No. 6465185
; GENERAL INFORMATION:
;  APPLICANT: Goldfine, Ira
;  APPLICANT:  Trischitta, Vincenzo
;  APPLICANT:  Pizzuti, Antonio
;  APPLICANT:  Vigneri, Riccardo
;  APPLICANT:  Frittitta, Lucia
;  TITLE OF INVENTION: Polymorphic Human PC-1 Sequences
;  FILE REFERENCE: 9076-089
;  CURRENT APPLICATION NUMBER: US/09/438,906
;  CURRENT FILING DATE:  1999-11-18
;  PRIOR APPLICATION NUMBER: 60/108,853
;  PRIOR FILING DATE: 1998-11-18
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
;   LENGTH: 873
;   TYPE: PRT
;   ORGANISM: H. sapiens
US-09-438-906-4
```

Query Match 51.9%; Score 2494.5; DB 4; Length 873;
Best Local Similarity 53.0%; Pred. No. 2e-234;
Matches 464; Conservative 138; Mismatches 250; Indels 23; Gaps 11;

```
Qy         10 EQPVKK----NTLK---KYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFDAS 62
              |:|::|     | |     ||: :||   ::   ||   ||      ::  ||: :||: :
Db          5 EEPLEKAARARTAKDPNTYKVLSLVLSVCVLTTILGCIFGLKPSCAKEVKSCKGRCFERT 64

Qy         63 FRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQK 122
              |    ||||| || : |:|| |:::||:|   || |||||||| ||  |||:|||||  :
Db         65 F---GNCRCDAACVELGNCCLDYQETCIEPEHIWTCNKFRCGEKRLTRSLCACSDDCKDQ 121

Qy        123 KDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMP 182
              ||| :| |||||| ||:|| |::  : ||| ||: || :|||:||||||||:||  |:|
Db        122 GDCCINYSSVCQGEKSWVEEPCESINEPQCPAGFETPPTLLFSLDGFRAEYLHTWGGLLP 181

Qy        183 NINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSK 242
              |:||| || ::| || :|||||||||||:|||||||||||||||| |||  :| :||| ||
Db        182 VISKLKKCGTYTKNMRPVYPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMNASFSLKSK 241

Qy        243 EQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTL 302
              |: || |: |:|:|:|| ||||| |:|||||:| ||| || ||  |||||||||||  :
Db        242 EKFNPEWYKGEPIWVTAKYQGLKSGTFFWPGSDVEINGIFPDIYKMYNGSVPFEERILAV 301

Qy        303 LKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLH 362
              |:|| ||| ||| |||:| |||||||: |||| |||||| ||   ||||:|||: |||
Db        302 LQWLQLPKDERPHFYTLYLEEPDSSGHSYGPVSSEVIKALQRVDGMVGMLMDGLKELNLH 361

Qy        363 NCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNSEE 422
              |:|:||::||||:|  | |  |:  |   :    :  ||| |:|  ::|  ::||| |
Db        362 RCLNLILISDHGMEQGSCKKYIYLNKYLGDVKNIKVIYGPAARLRPSDVPDKYYSFNYEG 421

Qy        423 IVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW-LAVRSKSNTNCGGG 481
              | ||||||:|:|||||||   ||||||:||: ||: :  ::| || ||:        || |
Db        422 IARNLSCREPNQHFKPYLKHFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSERKYCGSG 481

Qy        482 NHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNH 541
              || :| | :|:|:|: :|| ||   | : |||||||||||||| : ||||||||||||
Db        482 FHGSDNVFSNMQALFVGYGPGFKHGIEADTFENIEVYNLMCDLLNLTPAPNNGTHGSLNH 541

Qy        542 LLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQMLNLTQEEI 600
              ||| | | | | :||     | |     | ::| | | | :     :|      ||| |
Db        542 LLKNPVYTPKHPKEVHPLVQCPFTRN-PRDNLGCSCNPSI---LPIEDFQTQFNLTVAEE 597

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
                ||:|||||||     |||   :::||: : : ||:|:|||| :  : |
Db        598 KIIKHETLPYGRPRVLQKENTICLLSQHQFMSGYSQDILMPLWTSYTVDR--NDSFSTED 655

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQY-DALITSNLVPMYE 719
              :||  | |:| |     ||||  :  :::||| ||  |:  |   | :||:|:|:||||:
Db        656 FSNCLYQDFRIPLSPVHKCSFYKNNTKVSYGFLSPPQLNKNSSGIYSEALLTTNIVPMYQ 715

Qy        720 EFRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKH---LANTDVPI 776
              |: :| |||  || |:| |||||||||||:||::|||  |: : : :   : | :: |
Db        716 SFQVIWRYFHDTLLRKYAEERNGVNVVSGPVFDFDYDGRCDSLENLRQKRRVIRNQEILI 775

Qy        777 PTHYPVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHI 836
              |||:|:|||||||: | || :|   || | ||:|||  | |||  || :: ||||    |
Db        776 PTHFFIVLTSCKDTSQTPLHCEN-LDTLAFILPHRTDNSESCVHGKHDSSWVEELLMLHR 834

Qy        837 ARVRDVELLTGLDFYQDKVQPVSEILQLKTYLPTF 871
              ||: ||| :||| ||| : :|||:||:|||:||||
Db        835 ARITDVEHITGLSFYQQRKEPVSDILKLKTHLPTF 869
```

RESULT 8
US-08-346-455B-69
; Sequence 69, Application US/08346455B
; Patent No. 5731167
; GENERAL INFORMATION:
; APPLICANT: UNITED STATES OF AMERICA; DEPT.
; APPLICANT: OF HEALTH AND HUMAN SERVICES
; TITLE OF INVENTION: MOTILITY STIMULATING
; TITLE OF INVENTION: PROTEIN USEFUL IN CANCER DIAGNOSIS AND
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/346,455B
; FILING DATE: 28-NOV-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06613
; FILING DATE: 24-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/249,182
; FILING DATE: 25-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/822,043
; FILING DATE: 17-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36,434
; REFERENCE/DOCKET NUMBER: 2026-4149PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 915
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: Unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: No
; FEATURE:
; NAME/KEY: A2058 ATX protein
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
US-08-346-455B-69

Query Match 42.5%; Score 2040.5; DB 1; Length 915;
Best Local Similarity 40.5%; Pred. No. 5.2e-190;
Matches 379; Conservative 177; Mismatches 268; Indels 111; Gaps 11;

```
Qy         24 ACIVLLALLVIMSLGLGLGLGLRKLEK------------------QGSCRKKCFDASF 63
              :| ::      : : : ||    ::::                        |||: :||:
Db          9 SCQIISLFTFAVGVSICLGFTAHRIKRAEGWEEGPPTVLSDSPWTNISGSCKGRCFELQE 68

Qy         64 RGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQKK 123
                |  :|||  ||     || ||:: |::: | | | | |||| | | : | ||:||| :
Db         69 AGPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRCGEVRNEENACHCSEDCLARG 128

Qy        124 DCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMPN 183
              ||| :|: ||:||: |::::|:  : ::|| ||  ||:|:||:|||| |:      :|||
Db        129 DCCTNYQVVCKGESHWVDDDCEEIKAAECPAGFVRPPLIIFSVDGFRASYMKKGSKVMPN 188

Qy        184 INKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSKE 243
              | ||::|| || ||| :|||||||| ||: ||||||||||: |:|||   :  |  :|
Db        189 IEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDATFHLRGRE 248

Qy        244 QNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTLL 303
              : |  || |||:|:||  ||:|| |:||           |: :|:      | || |:|
Db        249 KFNHRWWGGQPLWITATKQGVKAGTFFW-----------SVVIPH------ERRILTIL 290
```

```
Qy        304 KWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSA-------------------------- 336
              :|| ||  |||  |  | |:|| |||  ||
Db        291 RWLTLPDHERPSVYAFYSEQPDFSGHKYGPFGPEESSYGSPFTPAKRPKRKVAPKRRQER 350

Qy        337 ------------------------RVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLA 371
                                       ::   |: :|    | ||:|||  |  |||:| :
Db        351 PVAPPKKRRRKIHRMDHYAAETRQDKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVG 410

Qy        372 DHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIR---AHNIPHDFFSFNSEEIVRNLS 428
              ||||:   |:: |::::|   ::   :  |   |||   ::|  :|      : |: ||:
Db        411 DHGMEDVTCDRTEFLSNYLTNVDDITLVPGTLGRIRSKFSNNAKYD-----PKAIIANLT 465

Qy        429 CRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW------LAVRSKSNTNC-GGG 481
              |:||||||||||   |||||||| | ||: :|| |:::|      | |  | :  |  |
Db        466 CKKPDQHFKPYLKQHLPKRLHYANNRRIEDIHLLVERRWHVARKPLDVYKKPSGKCFFQG 525

Qy        482 NHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNH 541
              :||::|:  ||: :|: :||:|| ||:| ||||||:||:|||||  ::|||||||||||||
Db        526 DHGFDNKVNSMQTVFVGYGPTFKYKTKVPPFENIELYNVMCDLLGLKPAPNNGTHGSLNH 585

Qy        542 LLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQMLNLTQEEI 600
              ||:   : |:  |||:: :  |          | | |   ::   :|:::|: |:
Db        586 LLRTNTFRPTMPEEVTRPNYPGIMYLQSDFDLGCTCDDKVEPKNKLDELNKRLHTK---- 641

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
               :| : :| :||| || :   : :||| :: ||: :   | :|:|||| :  : | :|
Db        642 GSTEERHLLYGRPAVLYR-TRYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDH 700

Qy        661 VPDCLRADVRVPPSESQKCSPYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              :  |:| |||| || || |  |  || :::|||:||  : : :::||| : :|:||||
Db        701 LTSCVRPDVRVSPSFSQNCLAYKNDKQMSYGFLFPPYLSSSPEAKYDAFLVTNMVPMYPA 760

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              |:::|:||  ||: |:|:|||||||:|||||||:|||  |  |:| :::  : :|:||||
Db        761 FKRVWNYFQRVLVKKYASERNGVNVISGPIFDYDYDGLHDTEDKIKQYVEGSSIPVPTHY 820

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              : ::||| : :   : | | | |  ||:|||| | |||   : |: ||||    | ||||
Db        821 YSIITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVR 880

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              |:| || |||::   :   ||| ||||| |:|: |
Db        881 DIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESEI 915


RESULT 9
US-08-977-221-69
; Sequence 69, Application US/08977221
; Patent No. 6084069
;  GENERAL INFORMATION:
;    APPLICANT:  UNITED STATES OF AMERICA; DEPT.
;    APPLICANT:  OF HEALTH AND HUMAN SERVICES
;    TITLE OF INVENTION:  MOTILITY STIMULATING
;    TITLE OF INVENTION:  PROTEIN USEFUL IN CANCER DIAGNOSIS AND
;    TITLE OF INVENTION:  THERAPY
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  U.S.A.
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/977,221
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/346,455
;      FILING DATE:  28-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/249,182
;      FILING DATE:  25-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/822,043
;      FILING DATE:  17-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DOROTHY R. AUTH
;      REGISTRATION NUMBER:  36,434
;      REFERENCE/DOCKET NUMBER:  2026-4149US3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;  INFORMATION FOR SEQ ID NO:  69:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  915
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  Unknown
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  No
;    FEATURE:
;      NAME/KEY:  A2058 ATX protein
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-08-977-221-69

  Query Match             42.5%;  Score 2040.5;  DB 3;  Length 915;
  Best Local Similarity   40.5%;  Pred. No. 5.2e-190;
  Matches  379;  Conservative  177;  Mismatches  268;  Indels  111;  Gaps   11;

Qy         24 ACIVLLALLVIMSLGLGLGLGLRKLEK-------------------QGSCRKKCFDASF 63
              :| ::      : : : ||    ::::                        |||: :||:
Db          9 SCQIISLFTFAVGVSICLGFTAHRIKRAEGWEEGPPTVLSDSPWTNISGSCKGRCFELQE 68

Qy         64 RGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQKK 123
               |  :|||  ||    || ||:: |::: | | | | |||| | | : | ||:||| :
Db         69 AGPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRCGEVRNEENACHCSEDCLARG 128

Qy        124 DCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMPN 183
              ||| :|: ||:||: |::::|:  : ::|| ||  ||:|:||:||||| |:     :|||
Db        129 DCCTNYQVVCKGESHWVDDDCEEIKAAECPAGFVRPPLIIFSVDGFRASYMKKGSKVMPN 188

Qy        184 INKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSKE 243
              | ||::|| || ||| :|||||||| ||: |||||||||: |:|||    :  | |  :|
Db        189 IEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDATFHLRGRE 248

Qy        244 QNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTLL 303
              : |  || |||:|:||  ||:|| |:||          |: :|:      | || |:|
Db        249 KFNHRWWGGQPLWITATKQGVKAGTFFW----------SVVIPH------ERRILTIL 290

Qy        304 KWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSA-------------------------- 336
              :|| ||  |||  |  | |:|| |||  ||
Db        291 RWLTLPDHERPSVYAFYSEQPDFSGHKYGPFGPEESSYGSPFTPAKRPKRKVAPKRRQER 350

Qy        337 ------------------------RVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLA 371
                                       ::   |: :|    | ||:|||  |  |||:| :
Db        351 PVAPPKKRRRKIHRMDHYAAETRQDKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVG 410

Qy        372 DHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIR---AHNIPHDFFSFNSEEIVRNLS 428
              ||||:   |:: |::::|   ::   :  |   |||   ::|  :|      : |: ||:
Db        411 DHGMEDVTCDRTEFLSNYLTNVDDITLVPGTLGRIRSKFSNNAKYD-----PKAIIANLT 465

Qy        429 CRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW------LAVRSKSNTNC-GGG 481
              |:||||||||||   |||||||| | ||: :|| |:::|      | |  | :  |  |
Db        466 CKKPDQHFKPYLKQHLPKRLHYANNRRIEDIHLLVERRWHVARKPLDVYKKPSGKCFFQG 525

Qy        482 NHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNH 541
              :||::|:  ||: :|: :||:|| ||:| ||||||:||:|||||  ::|||||||||||||
```

```
Db        526 DHGFDNKVNSMQTVFVGYGPTFKYKTKVPPFENIELYNVMCDLLGLKPAPNNGTHGSLNH 585

Qy        542 LLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQMLNLTQEEI 600

Db        586 LLRTNTFRPTMPEEVTRPNYPGIMYLQSDFDLGCTCDDKVEPKNKLDELNKRLHTK---- 641

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Db        642 GSTEERHLLYGRPAVLYR-TRYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDH 700

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Db        701 LTSCVRPDVRVSPSFSQNCLAYKNDKQMSYGFLFPPYLSSSPEAKYDAFLVTNMVPMYPA 760

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Db        761 FKRVWNYFQRVLVKKYASERNGVNVISGPIFDYDYDGLHDTEDKIKQYVEGSSIPVPTHY 820

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Db        821 YSIITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVR 880

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875

Db        881 DIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESEI 915


RESULT 10
US-09-483-831B-69
; Sequence 69, Application US/09483831B
; Patent No. 6417338
; GENERAL INFORMATION:
;  APPLICANT: STRACKE, MARY
;  APPLICANT:  LIOTTA, LANCE
;  APPLICANT:  SCHIFFMANN, ELLIOTT
;  APPLICANT:  KRUTZCH, HENRY
;  APPLICANT:  MURATA, JUN
;  TITLE OF INVENTION: AUTOTAXIN: MOTILITY STIMULATING PROTEIN USEFUL IN
;  TITLE OF INVENTION:  CANCER DIAGNOSIS AND THERAPY
;  FILE REFERENCE: 2026-4149US4
;  CURRENT APPLICATION NUMBER: US/09/483,831B
;  CURRENT FILING DATE:  2000-01-17
;  PRIOR APPLICATION NUMBER: 07/822,043
;  PRIOR FILING DATE: 1992-01-17
;  PRIOR APPLICATION NUMBER: 08/249,182
;  PRIOR FILING DATE: 1994-05-25
;  PRIOR APPLICATION NUMBER: 08/346,455
;  PRIOR FILING DATE: 1994-11-28
;  PRIOR APPLICATION NUMBER: 08/977,221
;  PRIOR FILING DATE: 1997-11-24
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 69
;   LENGTH: 915
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
;   OTHER INFORMATION: Polypeptide
US-09-483-831B-69

  Query Match             42.5%;  Score 2040.5;  DB 4;  Length 915;
  Best Local Similarity   40.5%;  Pred. No. 5.2e-190;
  Matches  379;  Conservative  177;  Mismatches  268;  Indels  111;  Gaps   11;

Qy         24 ACIVLLALLVIMSLGLGLGLGLRKLEK-------------------QGSCRKKCFDASF 63

Db          9 SCQIISLFTFAVGVSICLGFTAHRIKRAEGWEEGPPTVLSDSPWTNISGSCKGRCFELQE 68

Qy         64 RGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQKK 123

Db         69 AGPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRCGEVRNEENACHCSEDCLARG 128

Qy        124 DCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMPN 183

Db        129 DCCTNYQVVCKGESHWVDDDCEEIKAAECPAGFVRPPLIIFSVDGFRASYMKKGSKVMPN 188

Qy        184 INKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSKE 243

Db        189 IEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDATFHLRGRE 248

Qy        244 QNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSPPSIYMPYNGSVPFEERISTLL 303

Db        249 KFNHRWWGGQPLWITATKQGVKAGTFFW------------SVVIPH------ERRILTIL 290

Qy        304 KWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSA-------------------------- 336

Db        291 RWLTLPDHERPSVYAFYSEQPDFSGHKYGPFGPEESSYGSPFTPAKRPKRKVAPKRRQER 350

Qy        337 -------------------------RVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLA 371

Db        351 PVAPPKKRRRKIHRMDHYAAETRQDKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVG 410

Qy        372 DHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIR---AHNIPHDFFSFNSEEIVRNLS 428

Db        411 DHGMEDVTCDRTEFLSNYLTNVDDITLVPGTLGRIRSKFSNNAKYD-----PKAIIANLT 465

Qy        429 CRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW------LAVRSKSNTNC-GGG 481

Db        466 CKKPDQHFKPYLKQHLPKRLHYANNRRIEDIHLLVERRWHVARKPLDVYKKPSGKCFFQG 525

Qy        482 NHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNH 541

Db        526 DHGFDNKVNSMQTVFVGYGPTFKYKTKVPPFENIELYNVMCDLLGLKPAPNNGTHGSLNH 585

Qy        542 LLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQMLNLTQEEI 600

Db        586 LLRTNTFRPTMPEEVTRPNYPGIMYLQSDFDLGCTCDDKVEPKNKLDELNKRLHTK---- 641

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Db        642 GSTEERHLLYGRPAVLYR-TRYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDH 700

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Db        701 LTSCVRPDVRVSPSFSQNCLAYKNDKQMSYGFLFPPYLSSSPEAKYDAFLVTNMVPMYPA 760

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Db        761 FKRVWNYFQRVLVKKYASERNGVNVISGPIFDYDYDGLHDTEDKIKQYVEGSSIPVPTHY 820

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Db        821 YSIITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVR 880

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875

Db        881 DIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESEI 915


RESULT 11
PCT-US95-06613-69
; Sequence 69, Application PC/TUS9506613
;  GENERAL INFORMATION:
;    APPLICANT:  STRACKE, MARY; LIOTTA, LANCE;
;    APPLICANT:  SCHIFFMANN, ELLIOTT; KRUTZSCH,
;    APPLICANT:  HENRY; MURATA, JUN
;    TITLE OF INVENTION:  MOTILITY STIMULATING
;    TITLE OF INVENTION:  PROTEIN USEFUL IN CANCER DIAGNOSIS AND
;    TITLE OF INVENTION:  THERAPY
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
```

```
;       COUNTRY:  U.S.A.
;       ZIP:  10154
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy Disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  WordPerfect 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/06613
;       FILING DATE:  24-MAY-1995
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/346,455
;       FILING DATE:  28-NOV-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/249,182
;       FILING DATE:  25-MAY-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/822,043
;       FILING DATE:  17-JAN-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  DOROTHY R. AUTH
;       REGISTRATION NUMBER:  36,434
;       REFERENCE/DOCKET NUMBER:  2026-4149US2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 758-4800
;       TELEFAX:  (212) 751-6849
;   INFORMATION FOR SEQ ID NO:  69:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  915
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  Unknown
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  No
;     FEATURE:
;       NAME/KEY:  A2058 ATX protein
;       LOCATION:
;       IDENTIFICATION METHOD:
;       OTHER INFORMATION:
PCT-US95-06613-69

  Query Match             42.5%;  Score 2040.5;  DB 5;  Length 915;
  Best Local Similarity   40.5%;  Pred. No. 5.2e-190;
  Matches  379;  Conservative  177;  Mismatches  268;  Indels  111;  Gaps   11;

Qy         24 ACIVLLALLVIMSLGLGLGLGLRKLEK----------------------QGSCRKKCFDASF 63
              :| ::      : : : ||    ::::                      |||: :||:
Db          9 SCQIISLFTFAVGVSICLGFTAHRIKRAEGWEEGPPTVLSDSPWTNISGSCKGRCFELQE 68

Qy         64 RGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQKK 123
                |  :|||| ||    || ||:: |::: | | | | |||| | | : | ||:||| :
Db         69 AGPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRCGEVRNEENACHCSEDCLARG 128

Qy        124 DCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMPN 183
              ||| :|: ||:||: |::::|:  : ::|| ||  ||:|:||:|||| |:      :|||
Db        129 DCCTNYQVVCKGESHWVDDDCEEIKAAECPAGFVRPPLIIFSVDGFRASYMKKGSKVMPN 188

Qy        184 INKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSKE 243
              | ||::|| || ||| :|||||||| ||: |||||||||: |:|||    :  | |  :|
Db        189 IEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDATFHLRGRE 248

Qy        244 QNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTLL 303
              : |  || |||:|:||  ||:|| |:||  .          |: :|:       | || |:|
Db        249 KFNHRWWGGQPLWITATKQGVKAGTFFW----------SVVIPH------ERRILTIL 290

Qy        304 KWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSA-------------------------- 336
              :|| ||  |||  |  | |:|| |||  ||
Db        291 RWLTLPDHERPSVYAFYSEQPDFSGHKYGPFGPEESSYGSPFTPAKRPKRKVAPKRRQER 350

Qy        337 ------------------------RVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLA 371
                                      ::   |: :|   | ||:|||| |  |||:| :
Db        351 PVAPPKKRRRKIHRMDHYAAETRQDKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVG 410

Qy        372 DHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIR---AHNIPHDFFSFNSEEIVRNLS 428
              ||||:   |:: |::::|   ::   :  |   |||   ::|  :|      : |: ||:
Db        411 DHGMEDVTCDRTEFLSNYLTNVDDITLVPGTLGRIRSKFSNNAKYD-----PKAIIANLT 465

Qy        429 CRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW------LAVRSKSNTNC-GGG 481
              |:|||||||||   |||||||| | ||: :|| |:::|       | |  | :  |   |
Db        466 CKKPDQHFKPYLKQHLPKRLHYANNRRIEDIHLLVERRWHVARKPLDVYKKPSGKCFFQG 525

Qy        482 NHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNH 541
              :||::|:  ||: :|: :||:|| ||:| ||||||:||:||||| ::|||||||||||||
Db        526 DHGFDNKVNSMQTVFVGYGPTFKYKTKVPPFENIELYNVMCDLLGLKPAPNNGTHGSLNH 585

Qy        542 LLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQMLNLTQEEI 600
              ||:   : |:  |||:: :  |       | | |   ::   :|:::|: |:
Db        586 LLRTNTFRPTMPEEVTRPNYPGIMYLQSDFDLGCTCDDKVEPKNKLDELNKRLHTK---- 641

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
               :| : :| :||| || :   : :||| :: ||: :   | :|:|||| :  : | :|
Db        642 GSTEERHLLYGRPAVLYR-TRYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDH 700

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              :  |:| |||| || || |  |  || :::|||:||  : : :::||| : :|:||||
Db        701 LTSCVRPDVRVSPSFSQNCLAYKNDKQMSYGFLFPPYLSSSPEAKYDAFLVTNMVPMYPA 760

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              |:::|:|| ||: |:|:|||||||:||||||||:|||  |  |:| :::  : :|:||||
Db        761 FKRVWNYFQRVLVKKYASERNGVNVISGPIFDYDYDGLHDTEDKIKQYVEGSSIPVPTHY 820

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              : ::||| : :   : | | | |  ||:|||| | |||   : |: |||| | ||||
Db        821 YSIITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVR 880

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              |:| || |||:   :   ||| ||||| |:|: |
Db        881 DIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESEI 915


RESULT 12
US-08-346-455B-34
; Sequence 34, Application US/08346455B
; Patent No. 5731167
;  GENERAL INFORMATION:
;    APPLICANT:  UNITED STATES OF AMERICA; DEPT.
;    APPLICANT:  OF HEALTH AND HUMAN SERVICES
;    TITLE OF INVENTION:  MOTILITY STIMULATING
;    TITLE OF INVENTION:  PROTEIN USEFUL IN CANCER DIAGNOSIS AND
;    TITLE OF INVENTION:  THERAPY
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  U.S.A.
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/346,455B
;      FILING DATE:  28-NOV-1994
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/06613
;      FILING DATE:  24-MAY-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/249,182
;      FILING DATE:  25-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/822,043
```

```
;       FILING DATE:  17-JAN-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  DOROTHY R. AUTH
;       REGISTRATION NUMBER:  36,434
;       REFERENCE/DOCKET NUMBER:  2026-4149PCT
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 758-4800
;       TELEFAX:  (212) 751-6849
;   INFORMATION FOR SEQ ID NO:  34:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  829
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  Unknown
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  No
;     ORIGINAL SOURCE:
;       ORGANISM:  Human
;       STRAIN:
;       INDIVIDUAL ISOLATE:
;       DEVELOPMENTAL STAGE:
;       HAPLOTYPE:
;       TISSUE TYPE:
;       CELL TYPE:  Melanoma
;       CELL LINE:  A2058
;       ORGANELLE:
;     FEATURE:
;       NAME/KEY:
;       LOCATION:
;       IDENTIFICATION METHOD:
;       OTHER INFORMATION:  Putative protein
;       OTHER INFORMATION:  sequence of A2058 Autotaxin
US-08-346-455B-34

  Query Match             40.9%;  Score 1964.5;  DB 1;  Length 829;
  Best Local Similarity   42.4%;  Pred. No. 1.2e-182;
  Matches  363;  Conservative  164;  Mismatches  239;  Indels   91;  Gaps   10;

Qy          82 CWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQKKDCCADYKSVCQGETSWLE 141
               | ||:: |::: | | | | |||| | | : | ||:||| : ||| :|: ||:||: |::
Db           1 CHDFDELCLKTARGWECTKDRCGEVRNEENACHCSEDCLARGDCCTNYQVVCKGESHWVD 60

Qy         142 ENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKYMRAMY 201
               ::|:  : ::|| ||   ||:|:||:||||| |:     :|||| ||::|| || ||| :|
Db          61 DDCEEIKAAECPAGFVRPPLIIFSVDGFRASYMKKGSKVMPNIEKLRSCGTHSPYMRPVY 120

Qy         202 PTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSKEQNNPAWWHGQPMWLTAMY 261
               ||||||| ||: ||||||||||: |:|||   :  | |  :|: |  || |||:|:||
Db         121 PTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDATFHLRGREKFNHRWWGGQPLWITATK 180

Qy         262 QGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTLLKWLDLPKAERPRFYTMYF 321
               ||:|| |:||             |: :|:      | || |:|:|| ||  |||  |  |
Db         181 QGVKAGTFFW-----------SVVIPH------ERRILTILRWLTLPDHERPSVYAFYS 222

Qy         322 EEPDSSGHAGGPVSA---------------------------------------------- 336
               |:|| |||  ||
Db         223 EQPDFSGHKYGPFGPEESSYGSPFTPAKRPKRKVAPKRRQERPVAPPKKRRRKIHRMDHY 282

Qy         337 -------RVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCNKMEYMTDY 389
                    ::   |: :|   | ||:|||| |  |||:| : ||||:   |:: |::::|
Db         283 AAETRQDKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCDRTEFLSNY 342

Qy         390 FPRINFFYMYEGPAPRIR---AHNIPHDFFSFNSEEIVRNLSCRKPDQHFKPYLTPDLPK 446
               ::   :  |   |||   ::|  :|      : |: ||:|:|||||||||   |||
Db         343 LTNVDDITLVPGTLGRIRSKFSNNAKYD-----PKAIIANLTCKKPDQHFKPYLKQHLPK 397

Qy         447 RLHYAKNVRIDKVHLFVDQQW------LAVRSKSNTNC-GGGNHGYNNEFRSMEAIFLAH 499
               ||||| | ||: :|| |:::|      | |  | :  |  |:||::|:  ||: :|: :
Db         398 RLHYANNRRIEDIHLLVERRWHVARKPLDVYKKPSGKCFFQGDHGFDNKVNSMQTVFVGY 457

Qy         500 GPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNHLLKVPFYEPSHAEEVSKF 559
               ||:|| ||:| |||||:||:||||| ::||||||||||||||:   : |:  |||::
Db         458 GPTFKYKTKVPPFENIELYNVMCDLLGLKPAPNNGTHGSLNHLLRTNTFRPTMPEEVTRP 517

Qy         560 SVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQMLNLTQEEITATVKVNLPFGRPRVLQK 618
               :  |       : | | |   ::   :|:::|: |:       :| : :| :||| || :
Db         518 NYPGIMYLQSDDDLGCTCDDKVEPKNKLDELNKRLHTK----GSTEERHLLYGRPAVLYR 573

Qy         619 NVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPTVPDCLRADVRVPPSESQK 678
               :|||  :: ||: :  | :|:|||| :  : | :|  :  |:| |||| || ||
Db         574 -TRYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRVSPSFSQN 632

Qy         679 CSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEEFRKMWDYFHSVLLIKHAT 738
               |  |  || :::|||:||  : : :::||| : :|:|||  |:::|:||  ||: |:|:
Db         633 CLAYKNDKQMSYGFLFPPYLSSSPEAKYDAFLVTNMVPMYPAFKRVWNYFQRVLVKKYAS 692

Qy         739 ERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHYFVVLTSCKNKSHTPENCP 798
               |||||||:|||||||:|||  |  |:| :::  : :|:||||: ::||| : :   : |
Db         693 ERNGVNVISGPIFDYDYDGLHDTEDKIKQYVEGSSIPVPTHYYSIITSCLDFTQPADKCD 752

Qy         799 GWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVRDVELLTGLDFYQDKVQPV 858
               | | |  ||:|||| | |||   : |: ||||    | |||||:| || |||::   :
Db         753 GPLSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSRSY 812

Qy         859 SEILQLKTYLPTFETTI 875
               ||| ||||| |:|: |
Db         813 PEILTLKTYLHTYESEI 829


RESULT 13
US-08-977-221-34
; Sequence 34, Application US/08977221
; Patent No. 6084069
;  GENERAL INFORMATION:
;    APPLICANT:  UNITED STATES OF AMERICA; DEPT.
;    APPLICANT:  OF HEALTH AND HUMAN SERVICES
;    TITLE OF INVENTION:  MOTILITY STIMULATING
;    TITLE OF INVENTION:  PROTEIN USEFUL IN CANCER DIAGNOSIS AND
;    TITLE OF INVENTION:  THERAPY
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  U.S.A.
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/977,221
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/346,455
;      FILING DATE:  28-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/249,182
;      FILING DATE:  25-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/822,043
;      FILING DATE:  17-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DOROTHY R. AUTH
;      REGISTRATION NUMBER:  36,434
;      REFERENCE/DOCKET NUMBER:  2026-4149US3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  829
;      TYPE:  amino acid
```

```
;       STRANDEDNESS:  single
;       TOPOLOGY:  Unknown
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  No
;     ORIGINAL SOURCE:
;       ORGANISM:  Human
;       STRAIN:
;       INDIVIDUAL ISOLATE:
;       DEVELOPMENTAL STAGE:
;       HAPLOTYPE:
;       TISSUE TYPE:
;       CELL TYPE:  Melanoma
;       CELL LINE:  A2058
;       ORGANELLE:
;     FEATURE:
;       NAME/KEY:
;       LOCATION:
;       IDENTIFICATION METHOD:
;       OTHER INFORMATION:  Putative protein
;       OTHER INFORMATION:  sequence of A2058 Autotaxin
US-08-977-221-34

  Query Match             40.9%;  Score 1964.5;  DB 3;  Length 829;
  Best Local Similarity   42.4%;  Pred. No. 1.2e-182;
  Matches  363;  Conservative  164;  Mismatches  239;  Indels   91;  Gaps   10;

Qy         82 CWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQKKDCCADYKSVCQGETSWLE 141
              | ||:: |::: | | | | |||| | | : | ||:||| : ||| :|: ||:||: |::
Db          1 CHDFDELCLKTARGWECTKDRCGEVRNEENACHCSEDCLARGDCCTNYQVVCKGESHWVD 60

Qy        142 ENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKYMRAMY 201
              ::|:  : ::|| ||  ||:|:||:||||| |:     :|||| ||::|| || ||| :|
Db         61 DDCEEIKAAECPAGFVRPPLIIFSVDGFRASYMKKGSKVMPNIEKLRSCGTHSPYMRPVY 120

Qy        202 PTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSKEQNNPAWWHGQPMWLTAMY 261
              ||||||| ||: |||||||||: |:||| :  | |  :|: |  || |||:|:||
Db        121 PTKTFPNNLYTLATGLYPESHGIVGNSMYDPVFDATFHLRGREKFNHRWWGGQPLWITATK 180

Qy        262 QGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTLLKWLDLPKAERPRFYTMYF 321
              ||:|| |:||           |: :|:      | || |:|:|| ||  |||  |  |
Db        181 QGVKAGTPFW----------SVVIPH------ERRILTILRWLTLPDHERPSVYAFYS 222

Qy        322 EEPDSSGHAGGPVSA---------------------------------------------- 336
              |:|| |||  ||
Db        223 EQPDFSGHKYGPFGPEESSYGSPFTPAKRPKRKVAPKRRQERPVAPPKKRRRKIHRMDHY 282

Qy        337 -------RVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCNKMEYMTDY 389
                    ::   |: :|   | ||:|||| |  |||:| : ||||:   |:: |::::|
Db        283 AAETRQDKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCDRTEFLSNY 342

Qy        390 FPRINFFYMYEGPAPRIR---AHNIPHDFFSFNSEEIVRNLSCRKPDQHFKPYLTPDLPK 446
                 ::   :  |   |||   ::|  :|      : |: ||:|:|||||||||   |||
Db        343 LTNVDDITLVPGTLGRIRSKFSNNAKYD-----PKAIIANLTCKKPDQHFKPYLKQHLPK 397

Qy        447 RLHYAKNVRIDKVHLFVDQQW------LAVRSKSNTNC-GGGNHGYNNEFRSMEAIFLAH 499
              |||||  | ||: :|| |:::|      | |  | :  |  |:||::|:  ||: :|: :
Db        398 RLHYANNRRIEDIHLLVERRWHVARKPLDVYKKPSGKCFFQGDHGFDNKVNSMQTVFVGY 457

Qy        500 GPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNHLLKVPFYEPSHAEEVSKF 559
              ||:|| ||:| |||||:||:||||| ::|||||||||||||||:   : |:  |||::
Db        458 GPTFKYKTKVPPFENIELYNVMCDLLGLKPAPNNGTHGSLNHLLRTNTFRPTMPEEVTRP 517

Qy        560 SVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQMLNLTQEEITATVKVNLPFGRPRVLQK 618
              :  |      : | | |   ::   :|:::|: |:       :| : :| :||| || :
Db        518 NYPGIMYLQSDDDLGCTCDDKVEPKNKLDELNKRLHTK----GSTEERHLLYGRPAVLYR 573

Qy        619 NVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPTVPDCLRADVRVPPSESQK 678
              : :||| :: ||: :  | :|:|||| : : | :|  :  |:| |||| || ||
Db        574 -TRYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRVSPSFSQN 632

Qy        679 CSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEEFRKMWDYFHSVLLIKHAT 738
              |  |  || :::|||:||  : : :::||| : :|:|||| |:::|:|| ||: |:|:
Db        633 CLAYKNDKQMSYGFLFPPYLSSSPEAKYDAFLVTNMVPMYPAFKRVWNYFQRVLVKKYAS 692

Qy        739 ERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHYFVVLTSCKNKSHTPENCP 798
              |||||||:||||||:|||  |  |:| :::  : :|:|||||: ::||| : :   : |
Db        693 ERNGVNVISGPIFDYDYDGLHDTEDKIKQYVEGSSIPVPTHYYSIITSCLDFTQPADKCD 752

Qy        799 GWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVRDVELLTGLDFYQDKVQPV 858
              | | |  ||:|||| | |||   : |: ||||     | |||||:| || |||::   :
Db        753 GPLSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSRSY 812

Qy        859 SEILQLKTYLPTFETTI 875
                ||| ||||| |:|: |
Db        813 PEILTLKTYLHTYESEI 829


RESULT 14
US-09-483-831B-34
; Sequence 34, Application US/09483831B
; Patent No. 6417338
; GENERAL INFORMATION:
;  APPLICANT: STRACKE, MARY
;  APPLICANT:  LIOTTA, LANCE
;  APPLICANT:  SCHIFFMANN, ELLIOTT
;  APPLICANT:  KRUTZCH, HENRY
;  APPLICANT:  MURATA, JUN
;  TITLE OF INVENTION: AUTOTAXIN: MOTILITY STIMULATING PROTEIN USEFUL IN
;  TITLE OF INVENTION:  CANCER DIAGNOSIS AND THERAPY
;  FILE REFERENCE: 2026-4149US4
;  CURRENT APPLICATION NUMBER: US/09/483,831B
;  CURRENT FILING DATE:  2000-01-17
;  PRIOR APPLICATION NUMBER: 07/822,043
;  PRIOR FILING DATE: 1992-01-17
;  PRIOR APPLICATION NUMBER: 08/249,182
;  PRIOR FILING DATE: 1994-05-25
;  PRIOR APPLICATION NUMBER: 08/346,455
;  PRIOR FILING DATE: 1994-11-28
;  PRIOR APPLICATION NUMBER: 08/977,221
;  PRIOR FILING DATE: 1997-11-24
;  NUMBER OF SEQ ID NOS: 70
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
;   LENGTH: 829
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Putative protein sequence of A2058 Autotoxin
US-09-483-831B-34

  Query Match             40.9%;  Score 1964.5;  DB 4;  Length 829;
  Best Local Similarity   42.4%;  Pred. No. 1.2e-182;
  Matches  363;  Conservative  164;  Mismatches  239;  Indels   91;  Gaps   10;

Qy         82 CWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQKKDCCADYKSVCQGETSWLE 141
              | ||:: |::: | | | | |||| | | : | ||:||| : ||| :|: ||:||: |::
Db          1 CHDFDELCLKTARGWECTKDRCGEVRNEENACHCSEDCLARGDCCTNYQVVCKGESHWVD 60

Qy        142 ENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKYMRAMY 201
              ::|:  : ::|| ||  ||:|:||:||||| |:     :|||| ||::|| || ||| :|
Db         61 DDCEEIKAAECPAGFVRPPLIIFSVDGFRASYMKKGSKVMPNIEKLRSCGTHSPYMRPVY 120

Qy        202 PTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSKEQNNPAWWHGQPMWLTAMY 261
              ||||||| ||: |||||||||: |:||| :  | |  :|: |  || |||:|:||
Db        121 PTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDATFHLRGREKFNHRWWGGQPLWITATK 180

Qy        262 QGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTLLKWLDLPKAERPRFYTMYF 321
              ||:|| |:||           |: :|:      | || |:|:|| ||  |||  |  |
Db        181 QGVKAGTPFW----------SVVIPH------ERRILTILRWLTLPDHERPSVYAFYS 222

Qy        322 EEPDSSGHAGGPVSA---------------------------------------------- 336
              |:|| |||  ||
Db        223 EQPDFSGHKYGPFGPEESSYGSPFTPAKRPKRKVAPKRRQERPVAPPKKRRRKIHRMDHY 282

Qy        337 -------RVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCNKMEYMTDY 389
                    ::   |: :|   | ||:|||| |  |||:| : ||||:   |:: |::::|
```

```
Db        283 AAETRQDKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCDRTEFLSNY 342

Qy        390 FPRINFFYMYEGPAPRIR---AHNIPHDFFSFNSEEIVRNLSCRKPDQHFKPYLTPDLPK 446

Db        343 LTNVDDITLVPGTLGRIRSKFSNNAKYD-----PKAIIANLTCKKPDQHFKPYLKQHLPK 397

Qy        447 RLHYAKNVRIDKVHLFVDQQW------LAVRSKSNTNC-GGGNHGYNNEFRSMEAIFLAH 499

Db        398 RLHYANNRRIEDIHLLVERRWHVARKPLDVYKKPSGKCFFQGDHGFDNKVNSMQTVFVGY 457

Qy        500 GPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNHLLKVPFYEPSHAEEVSKF 559

Db        458 GPTFKYKTKVPPFENIELYNVMCDLLGLKPAPNNGTHGSLNHLLRTNTFRPTMPEEVTRP 517

Qy        560 SVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQMLNLTQEEITATVKVNLPFGRPRVLQK 618

Db        518 NYPGIMYLQSDDDLGCTCDDKVEPKNKLDELNKRLHTK----GSTEERHLLYGRPAVLYR 573

Qy        619 NVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPTVPDCLRADVRVPPSESQK 678

Db        574 -TRYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRVSPSFSQN 632

Qy        679 CSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEEFRKMWDYFHSVLLIKHAT 738

Db        633 CLAYKNDKQMSYGFLFPPYLSSSPEAKYDAPLVTNMVPMYPAFKRVWNYFQRVLVKKYAS 692

Qy        739 ERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHYFVVLTSCKNKSHTPENCP 798

Db        693 ERNGVNVISGPIFDYDYDGLHDTEDKIKQYVEGSSIPVPTHYYSIITSCLDFTQPADKCD 752

Qy        799 GWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVRDVELLTGLDFYQDKVQPV 858

Db        753 GPLSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSRSY 812

Qy        859 SEILQLKTYLPTFETTI 875

Db        813 PEILTLKTYLHTYESEI 829
```

```
RESULT 15
PCT-US95-06613-34
; Sequence 34, Application PC/TUS9506613
;  GENERAL INFORMATION:
;    APPLICANT:  STRACKE, MARY; LIOTTA, LANCE;
;    APPLICANT:  SCHIFFMANN, ELLIOTT; KRUTZSCH,
;    APPLICANT:  HENRY; MURATA, JUN
;    TITLE OF INVENTION:  MOTILITY STIMULATING
;    TITLE OF INVENTION:  PROTEIN USEFUL IN CANCER DIAGNOSIS AND
;    TITLE OF INVENTION:  THERAPY
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  U.S.A.
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/06613
;      FILING DATE:  24-MAY-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/346,455
;      FILING DATE:  28-NOV-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/249,182
;      FILING DATE:  25-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/822,043
;      FILING DATE:  17-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DOROTHY R. AUTH
;      REGISTRATION NUMBER:  36,434
;      REFERENCE/DOCKET NUMBER:  2026-4149US2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;  INFORMATION FOR SEQ ID NO:  34:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  829
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  Unknown
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  No
;    ORIGINAL SOURCE:
;      ORGANISM:  Human
;      STRAIN:
;      INDIVIDUAL ISOLATE:
;      DEVELOPMENTAL STAGE:
;      HAPLOTYPE:
;      TISSUE TYPE:
;      CELL TYPE:  Melanoma
;      CELL LINE:  A2058
;      ORGANELLE:
;    FEATURE:
;      NAME/KEY:
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:  Putative protein
;      OTHER INFORMATION:  sequence of A2058 Autotaxin
PCT-US95-06613-34

  Query Match             40.9%;  Score 1964.5;  DB 5;  Length 829;
  Best Local Similarity   42.4%;  Pred. No. 1.2e-182;
  Matches  363;  Conservative  164;  Mismatches  239;  Indels   91;  Gaps   10;

Qy         82 CWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQKKDCCADYKSVCQGETSWLE 141

Db          1 CHDFDELCLKTARGWECTKDRCGEVRNEENACHCSEDCLARGDCCTNYQVVCKGESHWVD 60

Qy        142 ENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKYMRAMY 201

Db         61 DDCEEIKAAECPAGFVRPPLIIFSVDGFRASYMKKGSKVMPNIEKLRSCGTHSPYMRPVY 120

Qy        202 PTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSKEQNNPAWWHGQPMWLTAMY 261

Db        121 PTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDATFHLRGREKFNHRWWGGQPLWITATK 180

Qy        262 QGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTLLKWLDLPKAERPRFYTMYF 321

Db        181 QGVKAGTFFW------------SVVIPH------ERRILTILRWLTLPDHERPSVYAFYS 222

Qy        322 EEPDSSGHAGGPVSA---------------------------------------------- 336

Db        223 EQPDFSGHKYGPFGPEESSYGSPFTPAKRPKRKVAPKRRQERPVAPPKKRRRKIHRMDHY 282

Qy        337 -------RVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCNKMEYMTDY 389

Db        283 AAETRQDKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVGDHGMEDVTCDRTEFLSNY 342

Qy        390 FPRINFFYMYEGPAPRIR---AHNIPHDFFSFNSEEIVRNLSCRKPDQHFKPYLTPDLPK 446

Db        343 LTNVDDITLVPGTLGRIRSKFSNNAKYD-----PKAIIANLTCKKPDQHFKPYLKQHLPK 397

Qy        447 RLHYAKNVRIDKVHLFVDQQW------LAVRSKSNTNC-GGGNHGYNNEFRSMEAIFLAH 499

Db        398 RLHYANNRRIEDIHLLVERRWHVARKPLDVYKKPSGKCFFQGDHGFDNKVNSMQTVFVGY 457

Qy        500 GPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNHLLKVPFYEPSHAEEVSKF 559
```

```
Db        458 GPTFKYKTKVPPFENIELYNVMCDLLGLKPAPNNGTHGSLNHLLRTNTFRPTMPEEVTRP 517

Qy        560 SVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQMLNLTQEEITATVKVNLPFGRPRVLQK 618

Db        518 NYPGIMYLQSDDDLGCTCDDKVEPKNKLDELNKRLHTK----GSTEERHLLYGRPAVLYR 573

Qy        619 NVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPTVPDCLRADVRVPPSESQK 678

Db        574 -TRYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDHLTSCVRPDVRVSPSFSQN 632

Qy        679 CSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEEFRKMWDYFHSVLLIKHAT 738

Db        633 CLAYKNDKQMSYGFLFPPYLSSSPEAKYDAFLVTNMVPMYPAFKRVWNYFQRVLVKKYAS 692

Qy        739 ERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHYFVVLTSCKNKSHTPENCP 798

Db        693 ERNGVNVISGPIFDYDYDGLHDTEDKIKQYVEGSSIPVPTHYYSIITSCLDFTQPADKCD 752

Qy        799 GWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVRDVELLTGLDFYQDKVQPV 858

Db        753 GPLSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVRDIEHLTSLDFFRKTSRSY 812

Qy        859 SEILQLKTYLPTFETTI 875

Db        813 PEILTLKTYLHTYESEI 829
```

Search completed: July 6, 2004, 13:21:04
Job time : 27 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 6, 2004, 13:19:53 ; Search time 56 Seconds
(without alignments)
4863.810 Million cell updates/sec

Title: US-10-005-480A-743
Perfect score: 4804
Sequence: 1 MESTLTLATEQPVKKNTLKK.....QPVSEILQLKTYLPTFETTI 875

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1276540 seqs, 311283816 residues

Total number of hits satisfying chosen parameters: 1276540

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4804	100.0	875	14	US-10-062-109A-743	Sequence 743, App
2	4804	100.0	875	14	US-10-062-109A-747	Sequence 747, App
3	4804	100.0	875	14	US-10-062-109A-751	Sequence 751, App
4	4804	100.0	875	14	US-10-005-480A-743	Sequence 743, App
5	4804	100.0	875	14	US-10-005-480A-747	Sequence 747, App
6	4804	100.0	875	14	US-10-005-480A-751	Sequence 751, App
7	4804	100.0	875	15	US-10-291-241-3	Sequence 3, Appli
8	4804	100.0	875	15	US-10-291-241-11	Sequence 11, Appl
9	4804	100.0	875	15	US-10-291-241-16	Sequence 16, Appl
10	4804	100.0	875	15	US-10-291-241-21	Sequence 21, Appl
11	4804	100.0	875	15	US-10-291-241-22	Sequence 22, Appl
12	4804	100.0	875	15	US-10-291-241-81	Sequence 81, Appl
13	4804	100.0	875	15	US-10-291-241-94	Sequence 94, Appl
14	4804	100.0	875	15	US-10-291-241-95	Sequence 95, Appl
15	4804	100.0	875	15	US-10-291-241-96	Sequence 96, Appl
16	4804	100.0	875	15	US-10-291-241-103	Sequence 103, App
17	4801	99.9	875	14	US-10-062-109A-745	Sequence 745, App
18	4801	99.9	875	14	US-10-062-109A-748	Sequence 748, App
19	4801	99.9	875	14	US-10-005-480A-745	Sequence 745, App
20	4801	99.9	875	14	US-10-005-480A-748	Sequence 748, App
21	4801	99.9	875	15	US-10-291-241-5	Sequence 5, Appli
22	4801	99.9	875	15	US-10-291-241-17	Sequence 17, Appl
23	4798	99.9	875	15	US-10-291-241-9	Sequence 9, Appli
24	4798	99.9	875	15	US-10-291-241-19	Sequence 19, Appl
25	4797	99.9	875	15	US-10-291-241-7	Sequence 7, Appli
26	4797	99.9	875	15	US-10-291-241-18	Sequence 18, Appl
27	4796	99.8	875	15	US-10-291-241-13	Sequence 13, Appl
28	4704	97.9	871	14	US-10-062-109A-765	Sequence 765, App
29	4704	97.9	871	14	US-10-005-480A-765	Sequence 765, App
30	4642	96.6	841	15	US-10-291-241-15	Sequence 15, Appl
31	4642	96.6	841	15	US-10-291-241-101	Sequence 101, App
32	4634	96.5	841	15	US-10-291-241-20	Sequence 20, Appl
33	4634	96.5	841	15	US-10-291-241-100	Sequence 100, App
34	4634	96.5	841	15	US-10-291-241-102	Sequence 102, App
35	2498.5	52.0	873	9	US-09-997-701-6	Sequence 6, Appli
36	2265	47.1	945	12	US-10-296-115-1102	Sequence 1102, Ap
37	2120	44.1	862	12	US-10-250-342-2	Sequence 2, Appli
38	2108	43.9	862	12	US-10-307-817-603	Sequence 603, App
39	2094.5	43.6	863	12	US-10-307-817-278	Sequence 278, App
40	2094.5	43.6	863	12	US-10-307-817-604	Sequence 604, App
41	2094.5	43.6	863	12	US-10-307-817-605	Sequence 605, App
42	2094.5	43.6	863	12	US-10-307-817-610	Sequence 610, App
43	2091.5	43.5	870	12	US-10-307-817-286	Sequence 286, App
44	2091.5	43.5	871	12	US-10-307-817-276	Sequence 276, App
45	2084	43.4	384	14	US-10-062-109A-749	Sequence 749, App

ALIGNMENTS

RESULT 1
US-10-062-109A-743
; Sequence 743, Application US/10062109A
; Publication No. US20030165505A1
; GENERAL INFORMATION:
; APPLICANT: Agensys
; APPLICANT: Challita-Eid, Pia M.
; APPLICANT: Raitano, Arthur B.
; APPLICANT: Faris, Mary
; APPLICANT: Hubert, Rene S.
; APPLICANT: Morrison, Karen Jane Meyrick
; APPLICANT: Jakobovits, Aya
; TITLE OF INVENTION: Nucleic Acid and Corresponding Protein
; TITLE OF INVENTION: Entitled 161P2F10B Useful in Treatment and Detection of
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 51158-20062.01
; CURRENT APPLICATION NUMBER: US/10/062,109A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 10/005,480
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 765
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 743
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-062-109A-743

Query Match 100.0%; Score 4804; DB 14; Length 875;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy 61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db        61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy       121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
pb       121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy       181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy       241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy       301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy       361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy       421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy       481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy       541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy       601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy       661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy       721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy       781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy       841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
            ||||||||||||||||||||||||||||||||||||
Db       841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


RESULT 2
US-10-062-109A-747
; Sequence 747, Application US/10062109A
; Publication No. US20030165505A1
; GENERAL INFORMATION:
;  APPLICANT: Agensys
;  APPLICANT:  Challita-Eid, Pia M.
;  APPLICANT:  Raitano, Arthur B.
;  APPLICANT:  Faris, Mary
;  APPLICANT:  Hubert, Rene S.
;  APPLICANT:  Morrison, Karen Jane Meyrick
;  APPLICANT:  Jakobovits, Aya
;  TITLE OF INVENTION: Nucleic Acid and Corresponding Protein
;  TITLE OF INVENTION:  Entitled 161P2F10B Useful in Treatment and Detection of
;  TITLE OF INVENTION:  Cancer
;  FILE REFERENCE: 51158-20062.01
;  CURRENT APPLICATION NUMBER: US/10/062,109A
;  CURRENT FILING DATE:  2002-01-31
;  PRIOR APPLICATION NUMBER: US 10/005,480
;  PRIOR FILING DATE: 2001-11-07
;  NUMBER OF SEQ ID NOS: 765
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 747
;   LENGTH: 875
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-062-109A-747

  Query Match          100.0%;  Score 4804;  DB 14;  Length 875;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy        61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy       121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy       181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy       241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy       301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy       361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy       421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy       481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy       541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy       601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy       661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy       721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy       781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy       841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
            ||||||||||||||||||||||||||||||||||||
Db       841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
```

```
RESULT 3
US-10-062-109A-751
; Sequence 751, Application US/10062109A
; Publication No. US20030165505A1
; GENERAL INFORMATION:
;  APPLICANT: Agensys
;  APPLICANT:  Challita-Eid, Pia M.
;  APPLICANT:  Raitano, Arthur B.
;  APPLICANT:  Faris, Mary
;  APPLICANT:  Hubert, Rene S.
;  APPLICANT:  Morrison, Karen Jane Meyrick
;  APPLICANT:  Jakobovits, Aya
;  TITLE OF INVENTION: Nucleic Acid and Corresponding Protein
;  TITLE OF INVENTION:  Entitled 161P2F10B Useful in Treatment and Detection of
;  TITLE OF INVENTION:  Cancer
;  FILE REFERENCE: 51158-20062.01
;  CURRENT APPLICATION NUMBER: US/10/062,109A
;  CURRENT FILING DATE:  2002-01-31
;  PRIOR APPLICATION NUMBER: US 10/005,480
;  PRIOR FILING DATE: 2001-11-07
;  NUMBER OF SEQ ID NOS: 765
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 751
;   LENGTH: 875
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-062-109A-751

  Query Match             100.0%;  Score 4804;  DB 14;  Length 875;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


RESULT 4
US-10-005-480A-743
; Sequence 743, Application US/10005480A
; Publication No. US20030191073A1
; GENERAL INFORMATION:
;  APPLICANT: Agensys
;  APPLICANT:  Challita-Eid, Pia M.
;  APPLICANT:  Raitano, Arthur B.
;  APPLICANT:  Faris, Mary
;  APPLICANT:  Hubert, Rene S.
;  APPLICANT:  Morrison, Karen Jane Meyrick
;  APPLICANT:  Jakobovits, Aya
;  TITLE OF INVENTION: Nucleic Acid and Corresponding Protein
;  TITLE OF INVENTION:  Entitled 161P2F10B Useful in Treatment and Detection of
;  TITLE OF INVENTION:  Cancer
;  FILE REFERENCE: 51158-20062.00
;  CURRENT APPLICATION NUMBER: US/10/005,480A
;  CURRENT FILING DATE:  2001-11-07
;  NUMBER OF SEQ ID NOS: 765
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 743
;   LENGTH: 875
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-005-480A-743

  Query Match             100.0%;  Score 4804;  DB 14;  Length 875;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
```

```
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875

Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


RESULT 5
US-10-005-480A-747
; Sequence 747, Application US/10005480A
; Publication No. US20030191073A1
; GENERAL INFORMATION:
;  APPLICANT: Agensys
;  APPLICANT:  Challita-Eid, Pia M.
;  APPLICANT:  Raitano, Arthur B.
;  APPLICANT:  Faris, Mary
;  APPLICANT:  Hubert, Rene S.
;  APPLICANT:  Morrison, Karen Jane Meyrick
;  APPLICANT:  Jakobovits, Aya
;  TITLE OF INVENTION: Nucleic Acid and Corresponding Protein
;  TITLE OF INVENTION:  Entitled 161P2F10B Useful in Treatment and Detection of
;  TITLE OF INVENTION:  Cancer
;  FILE REFERENCE: 51158-20062.00
;  CURRENT APPLICATION NUMBER: US/10/005,480A
;  CURRENT FILING DATE:  2001-11-07
;  NUMBER OF SEQ ID NOS: 765
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 747
;   LENGTH: 875
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-005-480A-747

  Query Match           100.0%;  Score 4804;  DB 14;  Length 875;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875

Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


RESULT 6
US-10-005-480A-751
; Sequence 751, Application US/10005480A
; Publication No. US20030191073A1
; GENERAL INFORMATION:
;  APPLICANT: Agensys
;  APPLICANT:  Challita-Eid, Pia M.
;  APPLICANT:  Raitano, Arthur B.
;  APPLICANT:  Faris, Mary
;  APPLICANT:  Hubert, Rene S.
;  APPLICANT:  Morrison, Karen Jane Meyrick
;  APPLICANT:  Jakobovits, Aya
;  TITLE OF INVENTION: Nucleic Acid and Corresponding Protein
;  TITLE OF INVENTION:  Entitled 161P2F10B Useful in Treatment and Detection of
;  TITLE OF INVENTION:  Cancer
```

```
;  FILE REFERENCE: 51158-20062.00
;  CURRENT APPLICATION NUMBER: US/10/005,480A
;  CURRENT FILING DATE:  2001-11-07
;  NUMBER OF SEQ ID NOS: 765
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 751
;   LENGTH: 875
;   TYPE: PRT
;   ORGANISM: Homo Sapien
US-10-005-480A-751

  Query Match             100.0%;  Score 4804;  DB 14;  Length 875;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


RESULT 7
US-10-291-241-3
; Sequence 3, Application US/10291241
; Publication No. US20030206905A1
; GENERAL INFORMATION:
;  APPLICANT: Agensys, Inc.
;  APPLICANT:  Aya Jakobovits
;  APPLICANT:  Arthur B. Raitano
;  APPLICANT:  Mary Faris
;  APPLICANT:  Rene S. Hubert
;  APPLICANT:  Wangmao Ge
;  APPLICANT:  Karen Jane Meyrick Morrison
;  APPLICANT:  Robert Kendall Morrison
;  APPLICANT:  Pia M. Challita-Eid
;  TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
;  TITLE OF INVENTION:  ENTITLED 161P2F10B USEFUL IN TREATMENT AND DETECTION OF
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 51158-20062.20
;  CURRENT APPLICATION NUMBER: US/10/291,241
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/005,480
;  PRIOR FILING DATE: 2001-11-07
;  PRIOR APPLICATION NUMBER: US 10/062,109
;  PRIOR FILING DATE: 2002-01-31
;  PRIOR APPLICATION NUMBER: US 60/282,739
;  PRIOR FILING DATE: 2001-04-10
;  NUMBER OF SEQ ID NOS: 103
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 875
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-291-241-3

  Query Match             100.0%;  Score 4804;  DB 15;  Length 875;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
```

```
Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
```

RESULT 8
US-10-291-241-11
; Sequence 11, Application US/10291241
; Publication No. US20030206905A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Aya Jakobovits
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Rene S. Hubert
; APPLICANT: Wangmao Ge
; APPLICANT: Karen Jane Meyrick Morrison
; APPLICANT: Robert Kendall Morrison
; APPLICANT: Pia M. Challita-Eid
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 161P2F10B USEFUL IN TREATMENT AND DETECTION OF
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 51158-20062.20
; CURRENT APPLICATION NUMBER: US/10/291,241
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 10/005,480
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 10/062,109
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: US 60/282,739
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 875
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-291-241-11

Query Match 100.0%; Score 4804; DB 15; Length 875;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 875; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
```

RESULT 9
US-10-291-241-16
; Sequence 16, Application US/10291241
; Publication No. US20030206905A1
; GENERAL INFORMATION:
; APPLICANT: Agensys, Inc.
; APPLICANT: Aya Jakobovits
; APPLICANT: Arthur B. Raitano
; APPLICANT: Mary Faris
; APPLICANT: Rene S. Hubert
; APPLICANT: Wangmao Ge
; APPLICANT: Karen Jane Meyrick Morrison
; APPLICANT: Robert Kendall Morrison
; APPLICANT: Pia M. Challita-Eid
; TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
; TITLE OF INVENTION: ENTITLED 161P2F10B USEFUL IN TREATMENT AND DETECTION OF
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 51158-20062.20

```
;  CURRENT APPLICATION NUMBER: US/10/291,241
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/005,480
;  PRIOR FILING DATE: 2001-11-07
;  PRIOR APPLICATION NUMBER: US 10/062,109
;  PRIOR FILING DATE: 2002-01-31
;  PRIOR APPLICATION NUMBER: US 60/282,739
;  PRIOR FILING DATE: 2001-04-10
;  NUMBER OF SEQ ID NOS: 103
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
;   LENGTH: 875
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-291-241-16

  Query Match             100.0%;  Score 4804;  DB 15;  Length 875;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


RESULT 10
US-10-291-241-21
; Sequence 21, Application US/10291241
; Publication No. US20030206905A1
; GENERAL INFORMATION:
;  APPLICANT: Agensys, Inc.
;  APPLICANT:  Aya Jakobovits
;  APPLICANT:  Arthur B. Raitano
;  APPLICANT:  Mary Faris
;  APPLICANT:  Rene S. Hubert
;  APPLICANT:  Wangmao Ge
;  APPLICANT:  Karen Jane Meyrick Morrison
;  APPLICANT:  Robert Kendall Morrison
;  APPLICANT:  Pia M. Challita-Eid
;  TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
;  TITLE OF INVENTION:  ENTITLED 161P2F10B USEFUL IN TREATMENT AND DETECTION OF
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 51158-20062.20
;  CURRENT APPLICATION NUMBER: US/10/291,241
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/005,480
;  PRIOR FILING DATE: 2001-11-07
;  PRIOR APPLICATION NUMBER: US 10/062,109
;  PRIOR FILING DATE: 2002-01-31
;  PRIOR APPLICATION NUMBER: US 60/282,739
;  PRIOR FILING DATE: 2001-04-10
;  NUMBER OF SEQ ID NOS: 103
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21
;   LENGTH: 875
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-291-241-21

  Query Match             100.0%;  Score 4804;  DB 15;  Length 875;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
```

```
Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


RESULT 11
US-10-291-241-22
; Sequence 22, Application US/10291241
; Publication No. US20030206905A1
; GENERAL INFORMATION:
;  APPLICANT: Agensys, Inc.
;  APPLICANT:  Aya Jakobovits
;  APPLICANT:  Arthur B. Raitano
;  APPLICANT:  Mary Faris
;  APPLICANT:  Rene S. Hubert
;  APPLICANT:  Wangmao Ge
;  APPLICANT:  Karen Jane Meyrick Morrison
;  APPLICANT:  Robert Kendall Morrison
;  APPLICANT:  Pia M. Challita-Eid
;  TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
;  TITLE OF INVENTION:  ENTITLED 161P2F10B USEFUL IN TREATMENT AND DETECTION OF
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 51158-20062.20
;  CURRENT APPLICATION NUMBER: US/10/291,241
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/005,480
;  PRIOR FILING DATE: 2001-11-07
;  PRIOR APPLICATION NUMBER: US 10/062,109
;  PRIOR FILING DATE: 2002-01-31
;  PRIOR APPLICATION NUMBER: US 60/282,739
;  PRIOR FILING DATE: 2001-04-10
;  NUMBER OF SEQ ID NOS: 103
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22
;   LENGTH: 875
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-291-241-22

  Query Match           100.0%;  Score 4804;  DB 15;  Length 875;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


RESULT 12
US-10-291-241-81
; Sequence 81, Application US/10291241
; Publication No. US20030206905A1
; GENERAL INFORMATION:
;  APPLICANT: Agensys, Inc.
;  APPLICANT:  Aya Jakobovits
;  APPLICANT:  Arthur B. Raitano
;  APPLICANT:  Mary Faris
;  APPLICANT:  Rene S. Hubert
;  APPLICANT:  Wangmao Ge
;  APPLICANT:  Karen Jane Meyrick Morrison
;  APPLICANT:  Robert Kendall Morrison
```

```
;  APPLICANT:  Pia M. Challita-Eid
;  TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
;  TITLE OF INVENTION:  ENTITLED 161P2F10B USEFUL IN TREATMENT AND DETECTION OF
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 51158-20062.20
;  CURRENT APPLICATION NUMBER: US/10/291,241
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/005,480
;  PRIOR FILING DATE: 2001-11-07
;  PRIOR APPLICATION NUMBER: US 10/062,109
;  PRIOR FILING DATE: 2002-01-31
;  PRIOR APPLICATION NUMBER: US 60/282,739
;  PRIOR FILING DATE: 2001-04-10
;  NUMBER OF SEQ ID NOS: 103
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 81
;   LENGTH: 875
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-291-241-81

  Query Match           100.0%;  Score 4804;  DB 15;  Length 875;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


RESULT 13
US-10-291-241-94
; Sequence 94, Application US/10291241
; Publication No. US20030206905A1
; GENERAL INFORMATION:
;  APPLICANT: Agensys, Inc.
;  APPLICANT:  Aya Jakobovits
;  APPLICANT:  Arthur B. Raitano
;  APPLICANT:  Mary Faris
;  APPLICANT:  Rene S. Hubert
;  APPLICANT:  Wangmao Ge
;  APPLICANT:  Karen Jane Meyrick Morrison
;  APPLICANT:  Robert Kendall Morrison
;  APPLICANT:  Pia M. Challita-Eid
;  TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
;  TITLE OF INVENTION:  ENTITLED 161P2F10B USEFUL IN TREATMENT AND DETECTION OF
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 51158-20062.20
;  CURRENT APPLICATION NUMBER: US/10/291,241
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/005,480
;  PRIOR FILING DATE: 2001-11-07
;  PRIOR APPLICATION NUMBER: US 10/062,109
;  PRIOR FILING DATE: 2002-01-31
;  PRIOR APPLICATION NUMBER: US 60/282,739
;  PRIOR FILING DATE: 2001-04-10
;  NUMBER OF SEQ ID NOS: 103
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 94
;   LENGTH: 875
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-291-241-94

  Query Match           100.0%;  Score 4804;  DB 15;  Length 875;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


RESULT 14
US-10-291-241-95
; Sequence 95, Application US/10291241
; Publication No. US20030206905A1
; GENERAL INFORMATION:
;  APPLICANT: Agensys, Inc.
;  APPLICANT:  Aya Jakobovits
;  APPLICANT:  Arthur B. Raitano
;  APPLICANT:  Mary Faris
;  APPLICANT:  Rene S. Hubert
;  APPLICANT:  Wangmao Ge
;  APPLICANT:  Karen Jane Meyrick Morrison
;  APPLICANT:  Robert Kendall Morrison
;  APPLICANT:  Pia M. Challita-Eid
;  TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
;  TITLE OF INVENTION:  ENTITLED 161P2F10B USEFUL IN TREATMENT AND DETECTION OF
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 51158-20062.20
;  CURRENT APPLICATION NUMBER: US/10/291,241
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/005,480
;  PRIOR FILING DATE: 2001-11-07
;  PRIOR APPLICATION NUMBER: US 10/062,109
;  PRIOR FILING DATE: 2002-01-31
;  PRIOR APPLICATION NUMBER: US 60/282,739
;  PRIOR FILING DATE: 2001-04-10
;  NUMBER OF SEQ ID NOS: 103
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 95
;   LENGTH: 875
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-291-241-95

  Query Match           100.0%;  Score 4804;  DB 15;  Length 875;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


RESULT 15
US-10-291-241-96
; Sequence 96, Application US/10291241
; Publication No. US20030206905A1
; GENERAL INFORMATION:
;  APPLICANT: Agensys, Inc.
;  APPLICANT:  Aya Jakobovits
;  APPLICANT:  Arthur B. Raitano
```

```
;  APPLICANT:  Mary Faris
;  APPLICANT:  Rene S. Hubert
;  APPLICANT:  Wangmao Ge
;  APPLICANT:  Karen Jane Meyrick Morrison
;  APPLICANT:  Robert Kendall Morrison
;  APPLICANT:  Pia M. Challita-Eid
;  TITLE OF INVENTION: NUCLEIC ACID AND CORRESPONDING PROTEIN
;  TITLE OF INVENTION:  ENTITLED 161P2F10B USEFUL IN TREATMENT AND DETECTION OF
;  TITLE OF INVENTION:  CANCER
;  FILE REFERENCE: 51158-20062.20
;  CURRENT APPLICATION NUMBER: US/10/291,241
;  CURRENT FILING DATE:  2003-03-18
;  PRIOR APPLICATION NUMBER: US 10/005,480
;  PRIOR FILING DATE: 2001-11-07
;  PRIOR APPLICATION NUMBER: US 10/062,109
;  PRIOR FILING DATE: 2002-01-31
;  PRIOR APPLICATION NUMBER: US 60/282,739
;  PRIOR FILING DATE: 2001-04-10
;  NUMBER OF SEQ ID NOS: 103
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 96
;   LENGTH: 875
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-291-241-96

  Query Match           100.0%;  Score 4804;  DB 15;  Length 875;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  875;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFD 60

Qy         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCL 120

Qy        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 QKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTL 180

Qy        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 MPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLS 240

Qy        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 SKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERIS 300

Qy        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRN 360

Qy        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNS 420

Qy        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 EEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGG 480

Qy        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLN 540

Qy        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 HLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEEI 600

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              ||||||||||||||||||||||||||||||||||||
Db        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875


Search completed: July  6, 2004, 13:26:07
Job time : 59 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: July 6, 2004, 13:11:52 ; Search time 26 Seconds
(without alignments)
3237.215 Million cell updates/sec

Title: US-10-005-480A-743
Perfect score: 4804
Sequence: 1 MESTLTLATEQPVKKNTLKK.....QPVSEILQLKTYLPTFETTI 875

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3972	82.7	875	1	A57080	cell surface antig
2	2498.5	52.0	925	1	A39216	nucleotide diphosp
3	2411.5	50.2	905	1	A27410	nucleotide diphosp
4	2043.5	42.5	885	1	A55453	nucleotide diphosp
5	2040.5	42.5	915	1	A55144	autotaxin precurso
6	743.5	15.5	496	2	T09931	probable phosphodi
7	737.5	15.4	479	2	T03293	probable phosphodi
8	712.5	14.8	457	2	T09932	probable phosphodi
9	680.5	14.2	461	2	T09933	probable phosphodi
10	644	13.4	477	1	A59390	probable phosphodi
11	637	13.3	485	2	T40657	probable phosphodi
12	632.5	13.2	477	1	A59391	probable phosphodi
13	632	13.2	300	2	A41179	protein kinase PC-
14	630.5	13.1	429	2	T33724	probable phosphodi
15	629.5	13.1	453	1	A59389	probable phosphodi
16	623	13.0	829	2	T19494	hypothetical prote
17	553	11.5	451	2	F87407	probable phosphodi
18	508	10.6	674	2	T19495	hypothetical prote
19	499.5	10.4	433	2	B82537	probable phosphodi
20	434	9.0	743	2	S19437	hypothetical prote
21	412.5	8.6	493	2	S50443	probable phosphodi
22	374	7.8	614	2	T30973	hypothetical prote
23	301	6.3	96	2	A25274	phosphodiesterase
24	222.5	4.6	434	2	F96958	aP superfamily [im
25	222.5	4.6	453	2	T16795	hypothetical prote
26	212	4.4	133	2	T09934	hypothetical prote
27	209	4.4	422	2	AB1252	pyrophosphatase ho
28	195.5	4.1	422	2	AH1614	weakly pyrophospha
29	150	3.1	360	2	T20867	hypothetical prote
30	146.5	3.0	370	2	C71052	hypothetical prote
31	138.5	2.9	369	2	A75091	phosphonoacetate h
32	135	2.8	476	2	JC7189	tubulointerstitial
33	132	2.7	474	2	A57480	tubulointerstitial
34	129	2.7	369	2	A34614	placental protein
35	126	2.6	2004	2	AC0314	probable membrane
36	121	2.5	4543	1	A53102	alpha-2-macroglobu
37	120.5	2.5	479	2	C75099	hypothetical prote
38	118	2.5	781	2	S43534	integrin beta3 - c
39	117	2.4	696	2	T27402	hypothetical prote
40	116.5	2.4	758	2	S62432	major facilitator
41	116	2.4	704	2	A34287	storage protein 2
42	116	2.4	2150	2	T32497	hypothetical prote
43	115.5	2.4	454	2	A46498	glucocorticoid-sen
44	115	2.4	397	2	T35880	hypothetical prote
45	114.5	2.4	1327	2	T14594	guanidine nucleoti

ALIGNMENTS

RESULT 1
A57080
cell surface antigen RB13-6 - rat
N;Contains: phosphodiesterase I (EC 3.1.4.1)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A57080
R;Deissler, H.; Lottspeich, F.; Rajewsky, M.F.
J. Biol. Chem. 270, 9849-9855, 1995
A;Title: Affinity purification and cDNA cloning of rat neural differentiation and tumor
A;Reference number: A57080; MUID:95247775; PMID:7730366
A;Accession: A57080
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-875 <DEI>
A;Cross-references: GB:Z47987; NID:g806378; PIDN:CAA88029.1; PID:g806379
C;Superfamily: nucleotide pyrophosphatase; somatomedin B homology
C;Keywords: exonuclease; glycoprotein; phosphoprotein; phosphoric diester hydrolase; su
F;1-22/Domain: cytosolic #status predicted <CYT>
F;23-45/Domain: transmembrane #status predicted <TMM>
F;46-875/Domain: extracellular #status predicted <EXT>
F;51-94/Domain: somatomedin B homology <SBH>
F;95-138/Domain: somatomedin B homology <SBH2>
F;206/Binding site: AMP (Thr) (covalent) #status predicted
F;237,280,289,533,574,594,702,789/Binding site: carbohydrate (Asn) (covalent) #status p

Query Match 82.7%; Score 3972; DB 1; Length 875;
Best Local Similarity 81.2%; Pred. No. 7.2e-272;
Matches 711; Conservative 77; Mismatches 86; Indels 2; Gaps 2;

```
Qy          1 MESTLTLATEQPVKKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQ-GSCRKKCF 59
              |:| | ||||:|:||::||:||| | ||||||||:|||||||||||| |:  ||||||||
Db          1 MDSRLALATEEPIKKDSLKRYKILCAVLLALLVIVSLGLGLGLGLRKPEEHIGSCRKKCF 60

Qy         60 DASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDC 119
              |:| |||| |||| | ||||||||||||||:||:|| || ||||||||||:||||:|||
Db         61 DSSHRGLEGCRCDSGCTDRGDCCWDFEDTCVKSTQIWTCNSFRCGETRLEAALCSCADDC 120

Qy        120 LQKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDT 179
              ||:|||| |||:|||||  |: | | ::|: ||||||| ||||||||||||||| || |
Db        121 LQRKDCCTDYKAVCQGEVPWVTEACASSQEPQCPEGFDQPPVILFSMDGFRAEYLQTWST 180

Qy        180 LMPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSL 239
              |:|||||||||:|||||||||||||||||||||||||||||||||||||||| |||||||
Db        181 LLPNINKLKTCGLHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVYLNKNFSL 240

Qy        240 SSKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERI 299
              || |::||||| |||:|||||||||||:|:| ||:||:|||||:|| |: |||:| ||
Db        241 SSVEKSNPAWWSGQPIWLTAMYQGLKAASYYWLGSDVAVNGSFPNIYRNYSNSVPYESRI 300

Qy        300 STLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQR 359
              :|||:|||||||||| |||:| |||||:|| ||||| |||||:|| ||||||||||||
```

```
Db        301 ATLLQWLDLPKAERPSFYTIYVEEPDSAGHKSGPVSAGVIKALQLVDDAFGMLMEGLKQR 360

Qy        360 NLHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFN 419
              |||||||||:|||||||| |:::|||||||| || ||||:|||||||  ||| |||:||
Db        361 NLHNCVNIIVLADHGMDQTSCDRVEYMTDYFPEIN-FYMYQGPAPRIRTRNIPQDFFTFN 419

Qy        420 SEEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCG 479
              ||||||:||||| |||||||||||||||||||||||||| ||:|||| |:| ::||
Db        420 SEEIVRDLSCRKSDQHFKPYLTPDLPKRLHYAKNVRIDKVHLMVDRQWLAYRNKGSSNCE 479

Qy        480 GGNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSL 539
              || |||||||:|||||||||||||||| :|||||||||||:|||| |||||||||:||||
Db        480 GGTHGYNNEFKSMEAIFLAHGPSFKEKTVIEPFENIEVYNLLCDLLHIQPAPNNGSHGSL 539

Qy        540 NHLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQEE 599
              ||||| |||:|||||||:|| : ||| ||| :||:| |  || | | ||||| |||:  |
Db        540 NHLLKAPFYQPSHAEELSKSAGCGFTTPLPKDSLNCSCLALQTSGQEEQVNQRLNLSGGE 599

Qy        600 ITATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPP 659
              ::|| | |||||||||:||| |||||||||||||||||:|||||||||: |||| |||
Db        600 VSATEKTNLPFGRPRVIQKNKDHCLLYHREYVSGFGKAMKMPMWSSYTVPKPGDTSSLPP 659

Qy        660 TVPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYE 719
              |||||||||||| ||||||||||||:|| |||||||    ::|||||||||||||:
Db        660 TVPDCLRADVRVDPSESQKCSFYLADQNIDHGFLYPPAIKGNNESQYDALITSNLVPMYK 719

Qy        720 EFRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTH 779
              ||:||||||| |||||:| |||||||||||||||||||||||||||| ::| ||||:|||
Db        720 EFKKMWDYFHKVLLIKYAIERNGVNVVSGPIFDYNYDGHFDAPDEITNYVAGTDVPVPTH 779

Qy        780 YFVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARV 839
              ||||||||||:|||::|||||||||::||||||||||| | | |||||| ||||||
Db        780 YFVVLTSCKNKTHTPDSCPGWLDVLPFVVPHRPTNVESCPENKAEDLWVEERFKAHIARV 839

Qy        840 RDVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
              |||||||||||||:| |||||||||||||||| |
Db        840 RDVELLTGLDFYQEKTQPVSEILQLKTYLPTFETII 875
```

RESULT 2
A39216
nucleotide diphosphatase (EC 3.6.1.9) - human
N;Contains: nucleotide pyrophosphatase (EC 3.6.1.9); phosphodiesterase I (EC 3.1.4.1)
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C;Accession: A39216; S21706; S23587; S51030
R;Buckley, M.F.; Loveland, K.A.; McKinstry, W.J.; Garson, O.M.; Goding, J.W.
J. Biol. Chem. 265, 17506-17511, 1990
A;Title: Plasma cell membrane glycoprotein PC-1. cDNA cloning of the human molecule, ami
A;Reference number: A39216; MUID:91009202; PMID:2211644
A;Accession: A39216
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-925 <BUC>
A;Cross-references: GB:J05654
R;Funakoshi, I.; Kato, H.; Horie, K.; Yano, T.; Hori, Y.; Kobayashi, H.; Inoue, T.; Suzu
Arch. Biochem. Biophys. 295, 180-187, 1992
A;Title: Molecular cloning of cDNAs for human fibroblast nucleotide pyrophosphatase.
A;Reference number: S21706; MUID:92246539; PMID:1315502
A;Accession: S21706
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-925 <FUN1>
A;Accession: S23587
A;Molecule type: protein
A;Residues: 116-121;247-271,'X',273-275;279-280,'X',282-283;303-316;362-364;449-465;482-

A;Note: it is uncertain whether Met-1 or Met-53 is the initiator
R;Belli, S.I.; Goding, J.W.
Eur. J. Biochem. 226, 433-443, 1994
A;Title: Biochemical characterization of human PC-1, an enzyme possessing alkaline phosp
A;Reference number: S51030; MUID:95094801; PMID:8001561
A;Accession: S51030
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-80 <BEL>
C;Genetics:
A;Gene: GDB:PDNP1; M6S1; NPPS
A;Cross-references: GDB:132615; OMIM:173335
A;Map position: 6q22-6q23
C;Superfamily: nucleotide pyrophosphatase; somatomedin B homology
C;Keywords: coenzyme A; glycoprotein; phosphoprotein; phosphoric diester hydrolase; tran
F;77-97/Domain: transmembrane #status predicted <TMM>
F;104-144/Domain: somatomedin B homology <SBH1>
F;145-188/Domain: somatomedin B homology <SBH2>
F;179,285,341,477,578,585,643,700,731,748/Binding site: carbohydrate (Asn) (covalent) #s
F;256/Binding site: AMP (Thr) (covalent) #status predicted

  Query Match 52.0%; Score 2498.5; DB 1; Length 925;
  Best Local Similarity 53.1%; Pred. No. 6e-168;
  Matches 465; Conservative 137; Mismatches 250; Indels 23; Gaps 11;

```
Qy         10 EQPVKK----NTLK---KYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCFDAS 62
              |:|::|     | |   ||: :||   ::   ||   ||     ::  ||: :||: :
Db         57 EEPLEKAARARTAKDPNTYKVLSLVLSVCVLTTILGCIFGLKPSCAKEVKSCKGRCFERT 116

Qy         63 FRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQK 122
              |    ||||| || : |:|| |:::||:|   || |||||||| ||  |||:|||||  |
Db        117 F---GNCRCDAACVELGNCCLDYQETCIEPEHIWTCNKFRCGEKRLTRSLCACSDDCKDK 173

Qy        123 KDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMP 182
              ||| :| |||||| ||:|| |::  : ||| ||: || :|||:|||||||:||  |:|
Db        174 GDCCINYSSVCQGEKSWVEEPCESINEPQCPAGFETPPTLLFSLDGFRAEYLHTWGGLLP 233

Qy        183 NINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSK 242
              |:||| || ::| || :|||||||||:|||||||||||||| ||| :| :||| ||
Db        234 VISKLKKCGTYTKNMRPVYPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMNASFSLKSK 293

Qy        243 EQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTL 302
              |: || |: |:|:|:|| |||||: |:|||||:| ||| || ||  |||||||||| :
Db        294 EKFNPEWYKGEPIWVTAKYQGLKSGTFFWPGSDVEINGIFPDIYKMYNGSVPFEERILAV 353

Qy        303 LKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLH 362
              |:|| ||| ||| ||:| |||||||: |||: |||||| ||   ||||:|||: |||
Db        354 LQWLQLPKDERPHFYTLYLEEPDSSGHSYGPVSSEVIKALQRVDGMVGMLMDGLKELNLH 413

Qy        363 NCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNSEE 422
              |:|:||::||||:|  |  |  |:  |    :  ||| |:|  ::|  ::||| |
Db        414 RCLNLILISDHGMEQGSCKKYIYLNKYLGDVKNIKVIYGPAARLRPSDVPDKYYSFNYEG 473

Qy        423 IVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW-LAVRSKSNTNCGGG 481
              | |||||:|:|||||||   ||||||:||: ||: :  ::| || ||:       || |
Db        474 IARNLSCREPNQHFKPYLKHFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSERKYCGSG 533

Qy        482 NHGYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNH 541
              || :| | :|:|:|: :|| ||   | : |||||||||||||| : ||||||||||||
Db        534 FHGSDNVFSNMQALFVGYGPGFKHGIEADTFENIEVYNLMCDLLNLTPAPNNGTHGSLNH 593

Qy        542 LLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQMLNLTQEEI 600
              ||| | | | | :||     | |    | ::| | | | :     :|       ||| |
Db        594 LLKNPVYTPKHPKEVHPLVQCPFTRN-PRDNLGCSCNPSI---LPIEDFQTQFNLTVAEE 649

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
                  ||:|||||||     |||   :::||: : : ||:|:|||| :  : |
Db        650 KIIKHETLPYGRPRVLQKENTICLLSQHQFMSGYSQDILMPLWTSYTVDR--NDSFSTED 707

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQY-DALITSNLVPMYE 719
              :||  | |:| |  |||||  :  :::||| ||  |: |   | :||:|:|:|||||:
Db        708 FSNCLYQDFRIPLSPVHKCSFYKNNTKVSYGFLSPPQLNKNSSGIYSEALLTTNIVPMYQ 767

Qy        720 EFRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKH---LANTDVPI 776
              |: :| ||| || |:| |||||||||||:|||::|||  |: : :  :  | :: |
Db        768 SFQVIWRYFHDTLLRKYAEERNGVNVVSGPVFDFDYDGRCDSLENLRQKRRVIRNQEILI 827

Qy        777 PTHYFVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHI 836
              |||:|:|||||||: | || :|   || | ||:|||  | |||  || :: ||||    |
```

```
Db        828 PTHFFIVLTSCKDTSQTPLHCEN-LDTLAFILPHRTDNSESCVHGKHDSSWVEELLMLHR 886

Qy        837 ARVRDVELLTGLDFYQDKVQPVSEILQLKTYLPTF 871
              ||: ||| :||| ||| : :|||:||:|||:|||| 
Db        887 ARITDVEHITGLSFYQQRKEPVSDILKLKTHLPTF 921


RESULT 3
A27410
nucleotide diphosphatase (EC 3.6.1.9) - mouse
N;Alternate names: nucleotide pyrophosphatase (EC 3.6.1.9); phosphodiesterase I (EC 3.1.
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C;Accession: A27410; I59055; S38354
R;van Driel, I.R.; Goding, J.W.
J. Biol. Chem. 262, 4882-4887, 1987
A;Title: Plasma cell membrane glycoprotein PC-1. Primary structure deduced from cDNA clc
A;Reference number: A27410; MUID:87165906; PMID:3104326
A;Accession: A27410
A;Molecule type: mRNA
A;Residues: 1-905 <VAN>
A;Cross-references: GB:J02700; NID:g200236; PIDN:AAA39893.1; PID:g200237
A;Note: the authors translated the codon CAG for residue 24 as Glu
R;van Driel, I.R.; Wilks, A.F.; Pietersz, G.A.; Goding, J.W.
Proc. Natl. Acad. Sci. U.S.A. 82, 8619-8623, 1985
A;Title: Murine plasma cell membrane antigen PC-1: Molecular cloning of cDNA and analysi
A;Reference number: I59055; MUID:86094275; PMID:3001713
A;Accession: I59055
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 203-219 <RES>
A;Cross-references: GB:M12552; NID:g200234; PIDN:AAA39892.1; PID:g200235
R;Belli, S.I.; van Driel, I.R.; Goding, J.W.
Eur. J. Biochem. 217, 421-428, 1993
A;Title: Identification and characterization of a soluble form of the plasma cell membra
A;Reference number: S38354; MUID:94039066; PMID:8223581
A;Accession: S38354
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 35-219 <BEL>
A;Cross-references: EMBL:L04516
C;Genetics:
A;Introns: 62/3; 87/1; 126/1; 168/1; 188/2
C;Superfamily: nucleotide pyrophosphatase; somatomedin B homology
C;Keywords: coenzyme A; glycoprotein; phosphoprotein; phosphoric diester hydrolase; tran
F;86-126/Domain: somatomedin B homology <SBH1>
F;127-170/Domain: somatomedin B homology <SBH2>
F;161,267,323,459,567,624/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;238/Binding site: AMP (Thr) (covalent) #status predicted

  Query Match            50.2%;  Score 2411.5;  DB 1;  Length 905;
  Best Local Similarity  50.7%;  Pred. No. 7.9e-162;
  Matches  445;  Conservative  152;  Mismatches  251;  Indels   29;  Gaps   10;

Qy         10 EQPVKK----------NTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEKQGSCRKKCF 59
              |:|::|           ||   ||:  :||   ::   ||   ||     ::  ||: :||
Db         39 EEPLEKADGARPAKDPNT---YKVLSLVLSVCVLTTILGCIFGLKPSCAKEVKSCKGRCF 95

Qy         60 DASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDC 119
              : :|    ||||| ||   |:|| ||::|||| | || |||||||| ||   :|||:|||
Db         96 ERTF---SNCRCDAACVSLGNCCLDFQETCVEPTHIWTCNKFRCGEKRLSRFVCSCADDC 152

Qy        120 LQKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDT 179
                ||| :| |||| : ||:|| |::    :||  |: || :|||:||||||||:||
Db        153 KTHNDCCINYSSVCQDKKSWVEETCESIDTPECPAEFESPPTLLFSLDGFRAEYLHTWGG 212

Qy        180 LMPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSL 239
              |:| |:||| || ::| || ||||||||||||:|||||||||||||||| |||  :| :|||
Db        213 LLPVISKLKNCGTYTKNMRPMYPTKTFPNHYSIVTGLYPESHGIIDNKMYDPKMNASFSL 272

Qy        240 SSKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERI 299
               |||: || |: |||:|:|| :| :|: |||||||:| |:|  | ||  ||||||||||
Db        273 KSKEKFNPLWYKGQPIWVTANHQEVKSGTYFWPGSDVEIDGILPDIYKVYNGSVPFEERI 332

Qy        300 STLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQR 359
              :|:|| ||  ||| ||||:| |||||||: ||||: |||||| ||   ||||:|||
Db        333 LAVLEWLQLPSHERPHFYTLYLEEPDSSGHSHGPVSSEVIKALQKVDRLVGMLMDGLKDL 392

Qy        360 NLHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFN 419
              |  |:|:||::||||:|  | |  |:  | :|   :  ||| |:|  ::|  ::|||
Db        393 GLDKCLNLILISDHGMEQGSCKKYVYLNKYLGDVNNVKVVYGPAARLRPTDVPETYYSFN 452

Qy        420 SEEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW-LAVRSKSNTNC 478
              | : :|||||:|:|||:||| | ||||||:||: ||: :  ::| || ||:        |
Db        453 YEALAKNLSCREPNQHFRPYLKPFLPKRLHFAKSDRIEPLTFYLDPQWQLALNPSERKYC 512

Qy        479 GGGNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPENIEVYNLMCDLLRIQPAPNNGTHGS 538
              | | || :| | :|:|:|: :||:||   ||: |||||||||||||| : ||||||:|||
Db        513 GSGFHGSDNLFSNMQALFIGYGPAFKHGAEVDSFENIEVYNLMCDLLGLIPAPNNGSHGS 572

Qy        539 LNHLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNLTQE 598
              |||||| | | ||| :|    | |    |: : | | |         ::   : |||| |
Db        573 LNHLLKKPIYNPSHPKEEGFLSQC----PIKSTSNDLGCTCDPWIVPIKDFEKQLNLTTE 628

Qy        599 EITATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLP 658
              :      : :|:| ||:| |   ||| :::::|:   : ||:|:|||   | :
Db        629 D-DDIYHMTVPYGGPRILLKQHHVCLLQQQQFLTGYSLDLLMPLWASYTF--LRNDQFSR 685

Qy        659 PTVPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQY-DALITSNLVPM 717
              :|:  |:|:| |   ||:|: ::  :::||| ||  || |:  | :||:|||:|||
Db        686 DDFSNCMYQDLRIPLSPVHKCSYYKSNSKLSYGFLTPPRLNRVSNHIYSEALLTSNIVPM 745

Qy        718 YEEFRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKH---LANTDV 774
              |: |: :| | |  || ::| |||||:||||||:||::||| :|: : : ::   : : ::
Db        746 YQSFQVIWHYLHDTLLQRYAHERNGINVVSGPVFDFDYDGRYDSLEILKQNSRVIRSQEI 805

Qy        775 PIPTHYFVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTA 834
              ||||:|:|||||  | ||  |   |:   :|:|||| |:|||  || |: |||| |
Db        806 LIPTHFFIVLTSCKQLSETPLECSA-LESSAYILPHRPDNIESCTHGKRESSWVEELLTL 864

Qy        835 HIARVRDVELLTGLDFYQDKVQPVSEILQLKTYLPTF 871
              | ||| ||||:||| ||||: : |||:|:|||:|| |
Db        865 HRARVTDVELITGLSFYQDRQESVSELLRLKTHLPIF 901


RESULT 4
A55453
nucleotide diphosphatase (EC 3.6.1.9) - rat
N;Contains: nucleotide pyrophosphatase (EC 3.6.1.9); phosphodiesterase I (EC 3.1.4.1)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
C;Accession: A55453; JU0187
R;Narita, M.; Goji, J.; Nakamura, H.; Sano, K.
J. Biol. Chem. 269, 28235-28242, 1994
A;Title: Molecular cloning, expression, and localization of a brain-specific phosphodies
A;Reference number: A55453; MUID:95050605; PMID:7961762
A;Accession: A55453
A;Molecule type: mRNA
A;Residues: 1-885 <NAR>
A;Cross-references: GB:D28560; NID:g464196; PIDN:BAA05910.1; PID:g464197
R;Narita, M.; Goji, J.; Sano, K.; Nakamura, H.
submitted to JIPID, February 1994
A;Description: Cloning and expression of brain-specific nucleotide diphosphohydrolase.
A;Reference number: JU0187
A;Accession: JU0187
A;Molecule type: mRNA
A;Residues: 1-66,'Q',68-81,'T',83-94,'C',96,'A',98-195,'A',197-514,'E',516-621,'E',623-6
A;Experimental source: strain Sprague-Dawley
C;Superfamily: nucleotide pyrophosphatase; somatomedin B homology
C;Keywords: coenzyme A; exonuclease; glycoprotein; phosphoprotein; phosphoric diester hy
F;54-97/Domain: somatomedin B homology <SBH1>
F;98-141/Domain: somatomedin B homology <SBH2>
F;53,150,396,408,522,608,829/Binding site: carbohydrate (Asn) (covalent) #status predict
F;207/Binding site: AMP (Thr) (covalent) #status predicted

  Query Match            42.5%;  Score 2043.5;  DB 1;  Length 885;
```

Best Local Similarity 41.7%; Pred. No. 6.8e-136;
Matches 380; Conservative 178; Mismatches 275; Indels 79; Gaps 12;

```
Qy         14 KKNTLKKYKIACIVLLALLVIMSLGLGLGLGLRKLEK------------------QGSC 54
Db          3 RQGCLGSFQVISLFTFAI----SVNICLGFTASRIKRAEWDEGPPTVLSDSPWTNTSGSC 58

Qy         55 RKKCFDASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCS 114
Db         59 KGRCFELQEVGPPDCRCDNLCKSYSSCCHDFDELCLKTVRGWECTKDRSGEVRNEENACH 118

Qy        115 CSDDCLQKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDL--PPVILFSMDGFRAE 172
Db        119 CPEDCLSRGDCCTNYQVVCKGESHWV----DDAARNQSSECLQVCPPPLIIFSVDGFRAS 174

Qy        173 YLYTWDTLMPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVN 232
Db        175 YMKKGSKVMPNIEKLRSCGTHVPYTRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPV 234

Qy        233 LNKNFSLSSKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGS 292
Db        235 FDASFHLRGREKFNHRWWGGQPLWITATKQGVRAGTFFW------------------SVS 276

Qy        293 VPFEERISTLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGML 352
Db        277 IPHERRILTILQWLSLPDNERPSVYAFYSEQPDFSGHKYGPFGPEMTNPLREIDKTVGQL 336

Qy        353 MEGLKQRNLHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIP 412
Db        337 MDGLKQLRLHRCVNVIFVGDHGMEDVTCDRTEFLSNYLTNVDDITLVPGTLGRIRAKSIN 396

Qy        413 HDFFSFNSEEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW----- 467
Db        397 NS--KYDPKTIIANLTCKKPDQHFKPYMKQHLPKRLHYANNRRIEDIHLLVDRRWHVARK 454

Qy        468 -LAVRSKSNTNC-GGGNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLL 525
Db        455 PLDVYKKPSGKCFFQGDHGFDNKVNSMQTVFVGYGPTFKYRTKVPPFENIELYNVMCDLL 514

Qy        526 RIQPAPNNGTHGSLNHLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFC-PHLQNST 584
Db        515 GLKPAPNNGTHGSLNHLLRTNTFRPTMPDEVSRPNYPGIMYLQSEFDLGCTCDDKVEPKN 574

Qy        585 QLEQVNQMLNL----------------TQEEITATVKV----NLPFGRPRVLQKNVDHC 623
Db        575 KLEELNKRLHTKGSTEAETGKFRGSKHENKKNLNGSVEPRKERHLLYGRPAVLYR-TSYD 633

Qy        624 LLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPTVPDCLRADVRVPPSESQKCSFYL 683
Db        634 ILYHTDFESGYSEIFLMPLWTSYTISKQAEVSSIPEHLTNCVRPDVRVSPGFSQNCLAYK 693

Qy        684 ADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEEFRKMWDYFHSVLLIKHATERNGV 743
Db        694 NDKQMSYGFLFPPYLSSSPEAKYDAFLVTNMVPMYPAFKRVWAYFQRVLVKKYASERNGV 753

Qy        744 NVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHYFVVLTSCKNKSHTPENCPGWLDV 803
Db        754 NVISGPIFDYNYDGLRDTEDEIKQYVEGSSIPVPTHYYSIITSCLDFTQPADKCDGPLSV 813

Qy        804 LPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVRDVELLTGLDFYQDKVQPVSEILQ 863
Db        814 SSFILPHRPDNDESCNSSEDESKWVEELMKMHTARVRDIEHLTGLDFYRKTSRSYSEILT 873

Qy        864 LKTYLPTFETTI 875
Db        874 LKTYLHTYESEI 885
```

RESULT 5
A55144
autotaxin precursor - human
N;Contains: phosphodiesterase I (EC 3.1.4.1)
C;Species: Homo sapiens (man)

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A55144; A42329
R;Murata, J.; Lee, H.Y.; Clair, T.; Krutzsch, H.C.; Arestad, A.A.; Sobel, M.E.; Liotta,
J. Biol. Chem. 269, 30479-30484, 1994
A;Title: cDNA cloning of the human tumor motility-stimulating protein, autotaxin, reveal
A;Reference number: A55144; MUID:95074054; PMID:7982964
A;Accession: A55144
A;Molecule type: mRNA
A;Residues: 1-915 <MUR>
A;Cross-references: GB:L35594; NID:g537905; PIDN:AAA64785.1; PID:g537906
A;Note: parts of this sequence were confirmed by peptide sequencing
R;Stracke, M.L.; Krutzsch, H.C.; Unsworth, E.J.; Arestad, A.; Cioce, V.; Schiffmann, E.;
J. Biol. Chem. 267, 2524-2529, 1992
A;Title: Identification, purification, and partial sequence analysis of autotaxin, a nov
A;Reference number: A42329; MUID:92129337; PMID:1733949
A;Accession: A42329
A;Molecule type: protein
A;Residues: 256-266;422-444;504-507,'AN';510,'X',511-515;533-548;'S',554-559,'N',561-562
A;Experimental source: A2058 melanoma cells
A;Note: sequence extracted from NCBI backbone (NCBIP:78526, NCBIP:78523, NCBIP:78521, NC
A;Note: a peptide fragment Tyr-Asp-Val-Pro-Trp-Asn-Glu-Thr-Ile was also found
C;Comment: This protein acts as an autocrine factor to stimulate tumor cell motility.
C;Genetics:
A;Gene: GDB:ATX
A;Cross-references: GDB:378346
A;Map position: 8q22-8qter
C;Superfamily: nucleotide pyrophosphatase; somatomedin B homology
C;Keywords: EF hand; glycoprotein; phosphoprotein; phosphoric diester hydrolase
F;55-98/Domain: somatomedin B homology <SBH1>
F;99-142/Domain: somatomedin B homology <SBH2>
F;54,463,577,859/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;210/Binding site: AMP (Thr) (covalent) #status predicted

Query Match 42.5%; Score 2040.5; DB 1; Length 915;
Best Local Similarity 40.5%; Pred. No. 1.2e-135;
Matches 379; Conservative 177; Mismatches 268; Indels 111; Gaps 11;

```
Qy         24 ACIVLLALLVIMSLGLGLGLGLRKLEK------------------QGSCRKKCFDASF 63
Db          9 SCQIISLFTFAVGVSICLGFTAHRIKRAEGWEEGPPTVLSDSPWTNISGSCKGRCFELQE 68

Qy         64 RGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCSDDCLQKK 123
Db         69 AGPPDCRCDNLCKSYTSCCHDFDELCLKTARGWECTKDRCGEVRNEENACHCSEDCLARG 128

Qy        124 DCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMPN 183
Db        129 DCCTNYQVVCKGESHWVDDDCEEIKAAECPAGFVRPPLIIFSVDGFRASYMKKGSKVMPN 188

Qy        184 INKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSKE 243
Db        189 IEKLRSCGTHSPYMRPVYPTKTFPNLYTLATGLYPESHGIVGNSMYDPVFDATFHLRGRE 248

Qy        244 QNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTLL 303
Db        249 KFNHRWWGGQPLWITATKQGVKAGTFFW------------SVVIPH------ERRILTIL 290

Qy        304 KWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSA-------------------------- 336
Db        291 RWLTLPDHERPSVYAFYSEQPDFSGHKYGPFGPEESSYGSPFTPAKRPKRKVAPKRRQER 350

Qy        337 ------------------------RVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLA 371
Db        351 PVAPPKKRRRKIHRMDHYAAETRQDKMTNPLREIDKIVGQLMDGLKQLKLRRCVNVIFVG 410

Qy        372 DHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIR---AHNIPHDFFSFNSEEIVRNLS 428
Db        411 DHGMEDVTCDRTEFLSNYLTNVDDITLVPGTLGRIRSKFSNNAKYD-----PKAIIANLT 465

Qy        429 CRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW------LAVRSKSNTNC-GGG 481
Db        466 CKKPDQHFKPYLKQHLPKRLHYANNRRIEDIHLLVERRWHVARKPLDVYKKPSGKCFFQG 525
```

```
Qy        482 NHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLNH 541
Db        526 DHGFDNKVNSMQTVFVGYGPTFKYKTKVPPFENIELYNVMCDLLGLKPAPNNGTHGSLNH 585

Qy        542 LLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFC-PHLQNSTQLEQVNQMLNLTQEEI 600
Db        586 LLRTNTFRPTMPEEVTRPNYPGIMYLQSDFDLGCTCDDKVEPKNKLDELNKRLHTK---- 641

Qy        601 TATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTSPLPPT 660
Db        642 GSTEERHLLYGRPAVLYR-TRYDILYHTDFESGYSEIFLMLLWTSYTVSKQAEVSSVPDH 700

Qy        661 VPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQYDALITSNLVPMYEE 720
Db        701 LTSCVRPDVRVSPSFSQNCLAYKNDKQMSYGFLFPPYLSSSPEAKYDAFLVTNMVPMYPA 760

Qy        721 FRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDVPIPTHY 780
Db        761 FKRVWNYFQRVLVKKYASERNGVNVISGPIFDYDYDGLHDTEDKIKQYVEGSSIPVPTHY 820

Qy        781 FVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTAHIARVR 840
Db        821 YSIITSCLDFTQPADKCDGPLSVSSFILPHRPDNEESCNSSEDESKWVEELMKMHTARVR 880

Qy        841 DVELLTGLDFYQDKVQPVSEILQLKTYLPTFETTI 875
Db        881 DIEHLTSLDFFRKTSRSYPEILTLKTYLHTYESEI 915
```

RESULT 6
T09931
probable phosphodiesterase I (EC 3.1.4.1) / nucleotide diphosphatase (EC 3.6.1.9) T16L4.
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Aug-2002
C;Accession: T09931
R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16897
A;Accession: T09931
A;Molecule type: DNA
A;Residues: 1-496 <BEV>
A;Cross-references: EMBL:AL079344; GSPDB:GN00062; ATSP:T16L4.190
A;Experimental source: cultivar Columbia; BAC clone T16L4
C;Genetics:
A;Gene: ATSP:T16L4.190
A;Map position: 4
C;Superfamily: human phosphodiesterase I / nucleotide pyrophosphatase 4
C;Keywords: coenzyme A; metalloprotein; phosphoric diester hydrolase

Query Match 15.5%; Score 743.5; DB 2; Length 496;
Best Local Similarity 41.0%; Pred. No. 1.4e-44;
Matches 159; Conservative 71; Mismatches 141; Indels 17; Gaps 9;

```
Qy        157 DLPPVILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKY-MRAMYPTKTFPNHYTIVTG 215
Db        105 DKPVVLLISSDGFR--FGYQFKTKLPSIHRLIANGTEAETGLIPVFPTLTFPNHYSIVTG 162

Qy        216 LYPESHGIIDNNMYDVNLNKNFSLSSKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSE 275
Db        163 LYPAYHGIINNHFVDPETGNVFTMASHE---PEWWLGEPLWETVVNQGLKAATYFWPGSE 219

Qy        276 VAINGSF---PSIYMPYNGSVPFEERISTLLKWLDLPKAERPRFYTMYFEEPDSSGHAGG 332
Db        220 VH-KGSWNCPQGLCQNYNGSVPFDDRVDTILSYFDLPSNEIPSFMTLYFEDPDHQGHQVG 278

Qy        333 PVSARVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCNKMEYMTDYFPR 392
Db        279 PDDPQITEAVVNIDRLIGRLIDGLEKRGIFEDVTMIMVGDHGMVGTCDKKLVVLDDLAPW 338

Qy        393 INF---FYMYEGPAPRIRAHNIPHDFFSFNSEEIVRNLSCRKPD--QHFKPYLTPDLPKR 447
Db        339 IKIPSSWVQYYTPLLAIQPPS-GHDAADIVA-KINEGLSSGKVENGKYLKVYLKEDLPSR 396
```

```
Qy        448 LHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGGGNHGYNNEFRSMEAIFLAHGPSFKEKT 507
Db        397 LHYVDSDRIPPIIGLVDEGFKVEQKKSKAKECGGAHGYDNAFFSMRTIFIGHGPMFSKGR 456

Qy        508 EVEPPFENIEVYNLMCDLLRIQPAPNNGT 535
Db        457 KVPSFENVQIYNVISSILGLKAAPNNGS 484
```

RESULT 7
T03293
probable phosphodiesterase I (EC 3.1.4.1) / nucleotide diphosphatase (EC 3.6.1.9) - rice
C;Species: Oryza sativa (rice)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Aug-2002
C;Accession: T03293
R;Hsing, Y.C.; Tsao, C.V.; Chow, T.; Hsieh, J.; Chen, Z.
submitted to the EMBL Data Library, April 1995
A;Description: Rice early embryogenesis gene.
A;Reference number: Z14889
A;Accession: T03293
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-479 <HSI>
A;Cross-references: EMBL:U25430; NID:g818848; PID:g818849
A;Experimental source: strain Tainung 67
C;Genetics:
A;Gene: OSE4
C;Superfamily: human phosphodiesterase I / nucleotide pyrophosphatase 4
C;Keywords: coenzyme A; metalloprotein; phosphoric diester hydrolase

Query Match 15.4%; Score 737.5; DB 2; Length 479;
Best Local Similarity 42.0%; Pred. No. 3.4e-44;
Matches 163; Conservative 64; Mismatches 140; Indels 21; Gaps 10;

```
Qy        159 PPVILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKY-MRAMYPTKTFPNHYTIVTGLY 217
Db         90 PVVILISSDGFRFGYQHKAAT--PHIHRLIGNGTSAATGLVPIFPTLTFPNHYSIATGLY 147

Qy        218 PESHGIIDNNMYDVNLNKNFSLSSKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVA 277
Db        148 PSSHGIINNYFPDPISGDYFTMSSHE---PKWWLGEPLWVTAADQGIQAATYFWPGSEVK 204

Qy        278 INGSF--PSIY-MPYNGSVPFEERISTLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPV 334
Db        205 -KGSWDCPDKYCRHYNGSVPFEERVDAILGYFDLPSDEMPQFLTLYFEDPDHQGHQVGPD 263

Qy        335 SARVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCNKMEYMTDYFPRIN 394
Db        264 DPAITEAVVRIDEMIGRLIAGLEERGVFEDVNVILVGDHGMVGTCDKKLVFLDELAPWIK 323

Qy        395 F---FYMYEGPAPRIRAHNIPHDFFSFNS--EEIVRNLSCRKPD--QHFKPYLTPDLPKR 447
Db        324 LEEDWVLSMTPLLAIR----PPDDMSLPDVVAKMNEGLGSGKVENGEYLRMYLKEDLPSR 379

Qy        448 LHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGGGNHGYNNEFRSMEAIFLAHGPSFKEKT 507
Db        380 LHYADSYRIPPIIGLPEEGYKVEMKRSDKNECGGAHGYDNAFFSMRTIFIAHGPRFEGGR 439

Qy        508 EVEPPFENIEVYNLMCDLLRIQPAPNNGT 535
Db        440 VVPSFENVEIYNVIASILNLEPAPNNGS 467
```

RESULT 8
T09932
probable phosphodiesterase I (EC 3.1.4.1) / nucleotide diphosphatase (EC 3.6.1.9) T16L4
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Aug-2002
C;Accession: T09932
R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16897
A;Accession: T09932

```
A;Molecule type: DNA
A;Residues: 1-457 <BEV>
A;Cross-references: EMBL:AL079344; GSPDB:GN00062; ATSP:T16L4.200
A;Experimental source: cultivar Columbia; BAC clone T16L4
C;Genetics:
A;Gene: ATSP:T16L4.200
A;Map position: 4
C;Superfamily: human phosphodiesterase I / nucleotide pyrophosphatase 4
C;Keywords: coenzyme A; metalloprotein; phosphoric diester hydrolase

  Query Match           14.8%;  Score 712.5;  DB 2;  Length 457;
  Best Local Similarity  38.5%;  Pred. No. 1.9e-42;
  Matches  166;  Conservative   76;  Mismatches  148;  Indels   41;  Gaps   12;

Qy      148 QQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKY-MRAMYPTKTF 206
Db       44 RRPQPPKKLNKPVVLLISCDGFRFGYQFKTET--PNIDLLISRGTEAKTGLIPVFPTMTF 101

Qy      207 PNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSKEQNNPAWWHGQPMWLTAMYQGLKA 266
Db      102 PNHYSIATGLYPASHGIIMNKFTDPVSGELFN----RNLNPKWWLGEPLWVTAVNQGLMA 157

Qy      267 ATYFWPGSEVAINGSF--PSIY--MPYNGSVPFEERISTLLKWLDLPKAERPRFYTMYFE 322
Db      158 ATYFWPGADVH-KGSWNCPKGFCKAPYNVSVPLEERVDTILNYFDLPEREIPDFMALYFD 216

Qy      323 EPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCNK 382
Db      217 EPDIQGHEYGPDDPRVTEAVSKVDKMIGRIIMGLEKRKVFSDVHVILLGDHGMVTNCDKK 276

Qy      383 MEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFN------------SEEIVRN--- 426
Db      277 VIYIDDLADWIKI------PADWIQDYS---PVLAMNPRWGKDVKNPGQKNAELVRKMNE 327

Qy      427 -LSCRK--PDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGGGNH 483
Db      328 ALSSGKVANGEFLQVYLKENLPQRLHYSDSSRIPPIIGMVGEGLMVKQNRTYVQECSGTH 387

Qy      484 GYNNEFRSMEAIFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGTHGSLNHLL 543
Db      388 GYDNMFFSMRSIFVGYGPRFRRGIKVPSFENVQVYNAVAEILGLRPAPNNGSSLFTRSLL 447

Qy      544 KVPFYEPSHAE 554
Db      448 -LPFGETSQVE 457


RESULT 9
T09933
probable phosphodiesterase I (EC 3.1.4.1) / nucleotide diphosphatase (EC 3.6.1.9) T16L4.
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Aug-2002
C;Accession: T09933
R;Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16897
A;Accession: T09933
A;Molecule type: DNA
A;Residues: 1-461 <BEV>
A;Cross-references: EMBL:AL079344; GSPDB:GN00062; ATSP:T16L4.210
A;Experimental source: cultivar Columbia; BAC clone T16L4
C;Genetics:
A;Gene: ATSP:T16L4.210
A;Map position: 4
C;Superfamily: human phosphodiesterase I / nucleotide pyrophosphatase 4
C;Keywords: coenzyme A; metalloprotein; phosphoric diester hydrolase

  Query Match           14.2%;  Score 680.5;  DB 2;  Length 461;
  Best Local Similarity  38.7%;  Pred. No. 3.4e-40;
  Matches  155;  Conservative   70;  Mismatches  137;  Indels   39;  Gaps   11;

Qy      159 PPVILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKY-MRAMYPTKTFPNHYTIVTGLY 217

Db       58 PVVLMISCDGFRFGYQFKTDT--PNIDLLISGGTEAKHGLIPVFPTMTFPNHYSIATGLY 115

Qy      218 PESHGIIDNNMYDVNLNKNFSLSSKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVA 277
Db      116 PAYHGIIMNKFTDPVTGEVFNKGLQ----PKWWLGEPLWVTAVNQGLKAVTYFWPGSEVL 171

Qy      278 ING-SFPSIYMP-YNGSVPFEERISTLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVS 335
Db      172 KSSWTCPEGYCPHFNLSVPLEERVDSVLSHFDHLEDEVPDLLMLYFDEPDQSGHNYGPDD 231

Qy      336 ARVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGM-DQTYCN-KMEYMTDYFPRI 393
Db      232 PRVTTAVSRVDKMIGRVIKGLKQREIFDEVHVILLGDHGMVTNCECNEKAIYIDDLADWI 291

Qy      394 NFFYMYEGPAPRIRAHNIPHDFFSFNSE------------EIVRNLSCR------KPDQ 434
Db      292 KI------PAAWIQAYS---PVLAINPQWGKDVENQSEKNAEVVAKMNEALSSGKVKNGE 342

Qy      435 HFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGGGNHGYNNEFRSMEA 494
Db      343 FLKVYLKEKLPERLHFSESYRIPPIIGIVGEGLMVRQNRTNAQVCYGDHGYDNELFSMRT 402

Qy      495 IFLAHGPSFKEKTEVEPFENIEVYNLMCDLLRIQPAPNNGT 535
Db      403 IFVGHGSRFSRGKKVPSFENVQIYNVVAELLGLRPAPNNGS 443


RESULT 10
A59390
probable phosphodiesterase I (EC 3.1.4.1) / nucleotide diphosphatase (EC 3.6.1.9) 5 - mo
N;Alternate names: ectonucleotide pyrophosphatase/phosphodiesterase 5; NPP5
C;Species: Mus musculus (house mouse)
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 16-Aug-2002
C;Accession: A59390
R;Gijsbers, R.; Ceulemans, H.; Stalmans, W.; Bollen, M.
J. Biol. Chem. 276, 1361-1368, 2001
A;Title: Structural and catalytic similarities between nucleotide pyrophosphatases/phosp
A;Reference number: A59390; MUID:21125673; PMID:11027689
A;Accession: A59390
A;Molecule type: DNA
A;Residues: 1-477 <GIJ>
A;Cross-references: GB:AF233377; NID:g12231525; PIDN:AAG49143.1
A;Experimental source: strain C57BL/6J
C;Superfamily: human phosphodiesterase I / nucleotide pyrophosphatase 4
C;Keywords: coenzyme A; metalloprotein; phosphoric diester hydrolase
F;72/Active site: Thr #status predicted

  Query Match           13.4%;  Score 644;  DB 1;  Length 477;
  Best Local Similarity  32.8%;  Pred. No. 1.3e-37;
  Matches  146;  Conservative   89;  Mismatches  164;  Indels   46;  Gaps   11;

Qy      144 CDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKYMRAMYPT 203
Db       16 CHSAPFSLQPEE---QKVLVVSFDGFRWDYLYKVPT--PHFHYIMKNGVHVNQVTNVFIT 70

Qy      204 KTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKNFSLSSKE-QNNPAWWHGQPMWLTAMYQ 262
Db       71 KTYPNHYTLVTGLFAENHGIVANDMFDPILNKSFSLEHMDIYDSKFWEEATPIWITNQRA 130

Qy      263 GLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFEERISTLLKWLDLPKAERP-RFYTMYF 321
Db      131 GHASGAAMWPGADVKIHDSFPTYYLPYNESVSFEDRVAKIIEWF---TAKDPINLGFLYW 187

Qy      322 EEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCN 381
Db      188 EEPDDTGHDVGPDSPLMGSVISDVDHKLGYLIKMLKRAKLWNNVNLIVTSDHGMTQCSKQ 247

Qy      382 KMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFFSFNSEEIVRNLSCRKPDQHFKPYLT 441
Db      248 RVIELDRYLDKEHYTLIDHSPVAAI----LPKE---GKFDEVYDALAGAHP--NLTVYKK 298

Qy      442 PDLPKRLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGGGNHGYNNEFRSMEAIFLAHGP 501
```

```
Db        299 EEIPERWHYKHNDRVQPIVAVADEGWYILQNKSD-DFLLGNHGYDNALAEMHPIFLAHGP 357

Qy        502 SFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLNHLLKVPFYEPSHAEEVSKFSV 561
              :|::    |   : ::|:|:| || :   |:||:  ::  ||
Db        358 AFRKNFTKEAMNSTDLYSLLCHLLNLTALPHNGSFWNVQDLL------------------ 399

Qy        562 CGFANPLPTESLDCFCPHLQNSTQL 586
                | | |         |: |::| |
Db        400 -SSATPKP-------IPYTQSTTLL 416


RESULT 11
T40657
probable phosphodiesterase I (EC 3.1.4.1) / nucleotide diphosphatase (EC 3.6.1.9) SPBC72
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 16-Aug-2002
C;Accession: T40657
R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.
submitted to the EMBL Data Library, November 1998
A;Reference number: Z21889
A;Accession: T40657
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-485 <LYN>
A;Cross-references: EMBL:AL034352; PIDN:CAA22177.1; GSPDB:GN00067; SPDB:SPBC725.05c
A;Experimental source: strain 972h-; cosmid c725
C;Genetics:
A;Gene: SPDB:SPBC725.05c
A;Map position: 2
C;Superfamily: human phosphodiesterase I / nucleotide pyrophosphatase 4
C;Keywords: coenzyme A; metalloprotein; phosphoric diester hydrolase

  Query Match             13.3%;  Score 637;  DB 2;  Length 485;
  Best Local Similarity   37.5%;  Pred. No. 4.2e-37;
  Matches  146;  Conservative   64;  Mismatches  165;  Indels   14;  Gaps    8;

Qy        161 VILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPES 220
              ||: |:|||||:|||    |  ||:  |    :|  ::   :|: ||||||||||||||||
Db         76 VIVISLDGFRADYLYRGFT--PNLLSLAERNVHVPFLIPSFPSITFPNHYTIVTGLYPES 133

Qy        221 HGIIDNNMYDVNLNKNFSLSSKEQN-NPAWW-HGQPMWLTAMYQGLKAATYFWPGSEVAI 278
              |||: || :|    | |  |  | | :| || |:|:|: |    :::| : |||:||
Db        134 HGIVSNNFFDPVTGKQFVNSMPECNKDPTWWDKGEPIWVNAERNNVRSAVHMWPGNEVEN 193

Qy        279 NGSFPSIYMPYNGSVPFEERISTLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARV 338
              :|  |:    :|      |:   :|:|||||  :||:    |    | ||| || |  :
Db        194 HGYRPTYSDGFNFDTTLREKKDRILEWLDLPDKDRPQLLLAYAPHVDMVGHAFGPDSPEL 253

Qy        339 IKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYM 398
              :| ||   | |:||||:||:   |||| |:|||  |  |:: :: : |      :
Db        254 NIIIQEVDIVIGELIEGLKKRNIDKHVNIIFLSDHGMAPTSDNRLIWLDNMFNLSAVAHR 313

Qy        399 YEGPAPRIRAHNIPHDFFSFNSEEIVRNLSCRKPD-QHFKPYLTPDLPKRLHYAKNVRID 457
              |    |  :   | : :  | :|       |  :::  |   |:| | ||: | ||
Db        314 DAWPLGGFRGESDLDDEYIY--ESLVNYSRSSLPSAENWNVYSKKDIPSRWHYSNNERIA 371

Qy        458 KVHLFVDQQWLAV----RSKSNTNCGGGNHGYNNEFRSMEAIFLAHGPSFK--EKTEVEP 511
              | :  |  |  |      |         | |||:|    | |:|:| | |||  :  :: |
Db        372 PVWMIPDVGWSLVSMLDHSPELEYEPLGVHGYDNLSPVMRALFIASGSSFKNFKGKKLAP 431

Qy        512 FENIEVYNLMCDLLRIQPAPNNGTH-GSL 539
              |:| |:| ::  :| :   |||||: |:|
Db        432 FQNTEIYGILSHILDLPAQPNNGTYEGAL 460


RESULT 12
A59391
probable phosphodiesterase I (EC 3.1.4.1) / nucleotide diphosphatase (EC 3.6.1.9) 5 - hu
N;Alternate names: ectonucleotide pyrophosphatase/phosphodiesterase 5; NPP5
C;Species: Homo sapiens (man)
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 16-Aug-2002
C;Accession: A59391
R;Smalley, C.
submitted to the EMBL Data Library, September 1999
A;Description: Contains actin, gamma pseudogene 9, a gene similar to PLASMA-CELL MEMBRAN
A;Reference number: A59391
A;Accession: A59391
A;Molecule type: DNA
A;Residues: 1-477 <SMA>
A;Cross-references: GB:AL035701; NID:g5924007; PIDN:CAB56566.1
R;Gijsbers, R.; Ceulemans, H.; Stalmans, W.; Bollen, M.
J. Biol. Chem. 276, 1361-1368, 2001
A;Title: Structural and catalytic similarities between nucleotide pyrophosphatases/phosp
A;Reference number: A59390; MUID:21125673; PMID:11027689
A;Contents: annotation
C;Genetics:
A;Gene: ENPP5
A;Map position: 6p11.2-6p21.1
C;Superfamily: human phosphodiesterase I / nucleotide pyrophosphatase 4
C;Keywords: coenzyme A; metalloprotein; phosphoric diester hydrolase
F;72/Active site: Thr #status predicted

  Query Match             13.2%;  Score 632.5;  DB 1;  Length 477;
  Best Local Similarity   34.4%;  Pred. No. 8.6e-37;
  Matches  140;  Conservative   80;  Mismatches  154;  Indels   33;  Gaps   10;

Qy        156 FDLPP----VILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKYMRAMYPTKTFPNHYT 211
              | | |    |:| | |||| :|||    |  |: : :   |:| | :  :: |||:|||||
Db         21 FSLQPDQQKVLLVSFDGFRWDYLYKVPT--PHFHYIMKYGVHVKQVTNVFITKTYPNHYT 78

Qy        212 IVTGLYPESHGIIDNNMYDVNLNKNFSLSSKE-QNNPAWWHGQPMWLTAMYQGLKAATYF 270
              :||||: |:|||: |:|:|   ||:|||      ::  |     |:|:|   |  :
Db         79 LVTGLFAENHGIVANDMFDPIRNKSFSLDHMNIYDSKFWEEATPIWITNQRAGHTSGAAM 138

Qy        271 WPGSEVAINGSFPSIYMPYNGSVPFEERISTLLKWLDLPKAERPRFYTMYFEEPDSSGHA 330
              |||::| |:  ||: |||| || ||:|:: :::|      |      :|:|:||  ||
Db        139 WPGTDVKIHKRFPTHYMPYNESVSFEDRVAKIIEW--FTSKEPINLGLLYWEDPDDMGHH 196

Qy        331 GGPVSARVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCNKMEYMTDYF 390
              || |  :   :  :|   | |:: ||:  | | :|:|: :|||| |    ::  :  |
Db        197 LGPDSPLMGPVISDIDKKLGYLIQMLKKAKLWNTLNLIITSDHGMTQCSEERLIELDQYL 256

Qy        391 PRINFFYMYEGPAPRIRAHNIPHDFFSFNSEEIVRNLSCRKPDQHFKPYLT----PDLPK 446
              : ::  : : |   |    :| :      :|:   |:       |  | ||     |:|:
Db        257 DKDHYTLIDQSPVAAI----LPKE---GKFDEVYEALT------HAHPNLTVYKKEDVPE 303

Qy        447 RLHYAKNVRIDKVHLFVDQQWLAVRSKSNTNCGGGNHGYNNEFRSMEAIFLAHGPSFKEK 506
              | ||  | ||  :    |: |  :::||: :  |||||:|    |  |||||||:|::
Db        304 RWHYKYNSRIQPIIAVADEGWHILQNKSD-DFLLGNHGYDNALADMHPIFLAHGPAFRKN 362

Qy        507 TEVEPPFENIEVYNLMCDLLRIQPAPNNGTHGSLNHLLK------VPF 547
                |   : ::| |:| || |   |:||:  ::  ||        ||:
Db        363 FSKEAMNSTDLYPLLCHLLNITAMPHNGSFWNVQDLLNSAMPRVVPY 409


RESULT 13
A41179
protein kinase PC-1 (EC 2.7.1.-) - bovine (fragments)
N;Alternate names: MAFP; major acidic fibroblast growth factor-stimulated phosphoprotei
C;Species: Bos primigenius taurus (cattle)
C;Date: 28-May-1992 #sequence_revision 22-Apr-1995 #text_change 30-Apr-1999
C;Accession: A41179; A49308
R;Oda, Y.; Kuo, M.D.; Huang, S.S.; Huang, J.S.
J. Biol. Chem. 266, 16791-16795, 1991
A;Title: The plasma cell membrane glycoprotein, PC-1, is a threonine-specific protein k
A;Reference number: A41179; MUID:91358477; PMID:1715869
A;Accession: A41179
A;Molecule type: protein
A;Residues: 1-26;36-56;59-67;68-133;134-144;145-267;268-300 <ODA>
A;Experimental source: liver
R;Oda, Y.; Kuo, M.D.; Huang, S.S.; Huang, J.S.
J. Biol. Chem. 268, 27318-27326, 1993
A;Title: The major acidic fibroblast growth factor (aFGF)-stimulated phosphoprotein fro
esterase activities.
A;Reference number: A49308; MUID:94086550; PMID:7505270
A;Accession: A49308
```

A;Molecule type: protein
A;Residues: 27-35,'X',37-58 <OD2>
A;Experimental source: liver
A;Note: sequence extracted from NCBI backbone (NCBIP:141583)
C;Superfamily: nucleotide pyrophosphatase; somatomedin B homology
C;Keywords: glycoprotein; phosphoprotein; phosphotransferase
F;1-25/Domain: somatomedin B homology (fragment) <SBH>
F;36/Binding site: AMP (Thr) (covalent) #status predicted

Query Match 13.2%; Score 632; DB 2; Length 300;
Best Local Similarity 21.8%; Pred. No. 4.9e-37;
Matches 174; Conservative 51; Mismatches 71; Indels 504; Gaps 13;

```
Qy         57 KCFDASFRGLENCRCDVACKDRGDCCWDFEDTCVESTRIWMCNKFRCGETRLEASLCSCS 116
              :||: :|    ||||| || : |:|| |
Db          2 RCFERTF---GNCRCDAACVELGNCCLD------------------------------ 26

Qy        117 DDCLQKKDCCADYKSVCQGETSWLEENCDTAQQSQCPEGFDLPPVILFSMDGFRAEYLYT 176

Db         27 ------------------------------------------------------------ 26

Qy        177 WDTLMPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPESHGIIDNNMYDVNLNKN 236
                          || :||||||||||:|||||||||||||||
Db         27 -------------------NMRPVYPTKTFPNHYSIVTGLYPESHGIIDNK--------- 58

Qy        237 FSLSSKEQNNPAWWHGQPMWLTAMYQGLKAATYFWPGSEVAINGSFPSIYMPYNGSVPFE 296
                        |:|:||||     |: |:|||||:| ||| || ||     |||||
Db         59 ---------------GEPIWLTA-----KSGTFFWPGSDVKINGIFPDIYXV---SVPFE 95

Qy        297 ERISTLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVSARVIKALQVVDHAFGMLMEGL 356
              |||  :|||| ||
Db         96 ERILAILKWLQLP------------------------------------------------ 108

Qy        357 KQRNLHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINFFYMYEGPAPRIRAHNIPHDFF 416
              |: ||| |:|:||::||||:|  |                                  ::
Db        109 KELNLHRCLNLILISDHGMEQGSCK--------------------------------YY 135

Qy        417 SFNSEEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAKNVRIDKVHLFVDQQW-LAVRSKSN 475
              ||: | |                          ||: ||:::  ::| || ||:
Db        136 SFDYEGI--------------------------AKSDRIERLTFYLDPQWQLALNPSER 168

Qy        476 TNCGGGNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPPFENIEVYNLMCDLLRIQPAPNNGT 535
              |||| || :| | :|:|:|          |||: |||||||||||||| : |||||||
Db        169 KYCGGGFHGSDNLFLNMQALFXXXXXXXXHSTEVDSFENIEVYNLMCDLLNLTPAPNNGT 228

Qy        536 HGSLNHLLKVPFYEPSHAEEVSKFSVCGFANPLPTESLDCFCPHLQNSTQLEQVNQMLNL 595
              |
Db        229 H------------------------------------------------------------ 229

Qy        596 TQEEITATVKVNLPFGRPRVLQKNVDHCLLYHREYVSGFGKAMRMPMWSSYTVPQLGDTS 655

Db        230 ------------------------------------------------------------ 229

Qy        656 PLPPTVPDCLRADVRVPPSESQKCSFYLADKNITHGFLYPPASNRTSDSQY-DALITSNL 714
                                          :::||| ||  :: |   | :||:|:|:
Db        230 -------------------------------LSYGFLSPPQLHKGSSQVYSEALLTTNI 257

Qy        715 VPMYEEFRKMWDYFHSVLLIKHATERNGVNVVSGPIFDYNYDGHFDAPDEITKHLANTDV 774
              ||||: |:
Db        258 VPMYQSFQ------------------------------------------------------ 265

Qy        775 PIPTHYFVVLTSCKNKSHTPENCPGWLDVLPFIIPHRPTNVESCPEGKPEALWVEERFTA 834
                                          : |            |:|||||
Db        266 -------------------------------VIH------------ESLWVEELLKL 279

Qy        835 HIARVRDVELLTGLDFYQDK 854
              | ||: ||| :||| ||| :
Db        280 HTARITDVEHITGLSFYQQR 299
```

RESULT 14
T33724
probable phosphodiesterase I (EC 3.1.4.1) / nucleotide diphosphatase (EC 3.6.1.9) - Zymo
C;Species: Zymomonas mobilis
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 16-Aug-2002
C;Accession: T33724
R;Lee, J.S.; Jin, S.J.; Kang, H.L.; Kang, H.S.
submitted to the EMBL Data Library, August 1998
A;Description: Sequence analysis of 67E10 cosmid clone of Zymomonas mobilis ZM4.
A;Reference number: Z21392
A;Accession: T33724
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-429 <LEE>
A;Cross-references: EMBL:AF086791; NID:g3820581; PID:g3089614; PIDN:AAC70363.1
C;Superfamily: human phosphodiesterase I / nucleotide pyrophosphatase 4
C;Keywords: coenzyme A; metalloprotein; phosphoric diester hydrolase

Query Match 13.1%; Score 630.5; DB 2; Length 429;
Best Local Similarity 35.4%; Pred. No. 1e-36;
Matches 135; Conservative 76; Mismatches 143; Indels 27; Gaps 10;

```
Qy        160 PVILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGLYPE 219
              |:|| |:|||||:|:     | ||:  |   | ::| |   :|: |||||||:||||||:
Db         49 PLILISIDGFRADYIKR--GLTPNLVSLAENGSYAKVMHPSFPSITFPNHYTLVTGLYPD 106

Qy        220 SHGIIDNNMYDVNL--NKNFSLSSKE-QNNPAWW-HGQPMWLTAMYQGLKAATYFWPGSE 275
              |||: ||| | ::  : :| :|  :  :  ||  |:|:|:||  ||: :|| ||||||
Db        107 HHGIVGNNMDDAHITPDSHFKMSDHQAATDRRWWDEGEPLWVTAEKQGVVSATMFWPGSE 166

Qy        276 VAINGSFPSIYMPYNGSVPFEERISTLLKWLDLPKAERPRFYTMYFEEPDSSGHAGGPVS 335
              | |  | ::  ::  ||| ||:  :  ||  || :||:| |:||  | :||  || |
Db        167 ADIEGVRPEMWRQFDSHVPFSERVDQVFSWLAYPKKKRPQFITLYFENVDHAGHLYGPDS 226

Qy        336 ARVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCNKMEYMTDYFPRINF 395
              |   | :|   | |::|||||:    |:::::|||| |  :::  :        :
Db        227 QEVNDNLVKIDQTIGQLVQGLKQRGIK--ANLVIVSDHGMAATSSDRVVALNKILDPSLY 284

Qy        396 FYM----YEGPAPRIRAHNIPHDFFSFNSEEIVRNLSCRKPDQHFKPYLTPDLPKRLHYA 451
              :    | |  |    |::      | | :              | : :    :| | ||
Db        285 HVVTGGAYAGIEPS-SGHSLKDLEPLFASHD-----------HMQCWPKQQIPARFHYG 331

Qy        452 KNVRIDKVHLFVDQQWLAVRSKS-NTNCGGGNHGYNNEFRSMEAIFLAHGPSFKEKTEVE 510
              :| |:  |    :  |  :   |   :   |||||:|:   | |:|:|:||:|::   :|
Db        332 QNPRVPAVVCAAEVGWSIMGDDSMAAHATKGNHGYDNQTPEMGALFIANGPAFQKHKVIE 391

Qy        511 PFENIEVYNLMCDLLRIQPAP 531
              :||:|  |:  :|::: ||
Db        392 SMDNIDVQPLVAQVLKLK-AP 411
```

RESULT 15
A59389
probable phosphodiesterase I (EC 3.1.4.1) / nucleotide diphosphatase (EC 3.6.1.9) 4 - hu
N;Alternate names: ectonucleotide pyrophosphatase/phosphodiesterase 4; KIAA0879 protein;
C;Species: Homo sapiens (man)
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 16-Aug-2002
C;Accession: A59389; B59391
R;Nagase, T.; Ishikawa, K.; Suyama, M.; Kikuno, R.; Hirosawa, M.; Miyajima, N.; Tanaka,
DNA Res. 5, 355-364, 1998
A;Title: Prediction of the coding sequences of unidentified human genes. XII. The comple
A;Reference number: A59389; MUID:99156230; PMID:10048485
A;Accession: A59389
A;Molecule type: mRNA
A;Residues: 1-453 <NAG>
A;Cross-references: GB:AB020686; NID:g4240247; PIDN:BAA74902.1
A;Experimental source: adult brain
R;Gijsbers, R.; Ceulemans, H.; Stalmans, W.; Bollen, M.
J. Biol. Chem. 276, 1361-1368, 2001
A;Title: Structural and catalytic similarities between nucleotide pyrophosphatases/phospl
A;Reference number: A59390; MUID:21125673; PMID:11027689
A;Contents: annotation
R;Smalley, C.
submitted to the EMBL Data Library, September 1999
A;Description: Contains actin, gamma pseudogene 9, a gene similar to PLASMA-CELL MEMBRAN
A;Reference number: A59391

```
A;Accession: B59391
A;Molecule type: DNA
A;Residues: 1-453 <SMA>
A;Cross-references: GB:AL035701; NID:g5924008; PIDN:CAB56567.1
C;Genetics:
A;Gene: ENPP4
A;Map position: 6p11.2-6p21.1
C;Superfamily: human phosphodiesterase I / nucleotide pyrophosphatase 4
C;Keywords: coenzyme A; metalloprotein; phosphoric diester hydrolase
F;70/Active site: Thr #status predicted

  Query Match             13.1%;  Score 629.5;  DB 1;  Length 453;
  Best Local Similarity   35.5%;  Pred. No. 1.3e-36;
  Matches  139;  Conservative   76;  Mismatches  153;  Indels   23;  Gaps   12;

Qy        158 LPP-VILFSMDGFRAEYLYTWDTLMPNINKLKTCGIHSKYMRAMYPTKTFPNHYTIVTGL 216
              ||| ::| | |||||:||  ::   |::      |:  :::: :: |||||||:|||||
Db         24 LPPKLLLVSFDGFRADYLKNYE--FPHLQNFIKEGVLVEHVKNVFITKTFPNHYSIVTGL 81

Qy        217 YPESHGIIDNNMYDVNLNKNFSLSSKEQNNPAWWH-GQPMWLT-AMYQGLKAATYFWPGS 274
              | |||||: |:|||     |:|| |:    :| ||:   |:|:|  : :   :|   |||: .
Db         82 YEESHGIVANSMYDAVTKKHFSDSN--DKDPFWWNEAVPIWVTNQLQENRSSAAAMWPGT 139

Qy        275 EVAINGSFPSIYMPYNGSVPFEERISTLLKWLDLPKAERP-RFYTMYFEEPDSSGHAGGP 333
              :| |: :  | :| || || ||||:: :  ||:   :  |  | |:|:||||:|||  ||
Db        140 DVPIHDTISSYFMNYNSSVSFEERLNNITMWLN--NSNPPVTFATLYWEEPDASGHKYGP 197

Qy        334 V-SARVIKALQVVDHAFGMLMEGLKQRNLHNCVNIILLADHGMDQTYCNKMEYMTDYFPR 392
                : : |: :|   | |:: ||   |   :|:|: :|||| |   :::  :
Db        198 EDKENMSRVLKKIDDLIGDLVQRLKMLGLWENLNVIITSDHGMTQCSQDRLINLDSCIDH 257

Qy        393 INFFYMYEGPAPRIRAHNIPHDFFSFNSEEIVRNLSCRKPDQHFKPYLTPDLPKRLHYAK 452
              :  :   |   |    :|      |  |:   |    |  |   ||  |:| | :|
Db        258 SYYTLIDLSPVAAI----LP----KINRTEVYNKLKNCSP--HMNVYLKEDIPNRFYYQH 307

Qy        453 NVRIDKVHLFVDQQWLAVRSKSNTNCGGGNHGYNNEFRSMEAIFLAHGPSFKEKTEVEPF 512
              | ||  : |  |: |  | ::|:     |:|||:|   ||     ||||:| :  :
Db        308 NDRIQPIILVADEGWTIVLNESSQKL--GDHGYDNSLPSMHPFLAAHGPAFHKGYKHSTI 365

Qy        513 ENIEVYNLMCDLLRIQPAPNNGTHGSLNHLL 543
              :::| :|| :| ::| ||||| |    ||
Db        366 NIVDIYPMMCHILGLKPHPNNGTFGHTKCLL 396


Search completed: July  6, 2004, 13:20:28
Job time : 29 secs
```